Feasibility of a GPSG French Grammar

Gabriel G. Bès, Karine Baschung

Feasibility of a GPSG French Grammar

Gabriel G. Bès and Karine Baschung
Formation doctorale Linguistique et informatique
Université Blaise-Pascal, Clermont-Ferrand

ESPRIT Project 393 ACORD, 1985

Abstract

The goal of this report is to study the feasibility of specifying a French grammar within the formalism proposed in GPSG.

The linguistic specification of this formalism is revised in part 1, "A review of GPSG". As we feel that one of the less dangerous ways to evaluate a linguistic model is to try to specify particular grammars, a variety of French constructions were studied and specified with some details along the lines of the GPSG formalism; the condensed results are presented in part 2, "A GPSG French Grammar" and in part 3, "Some remarks on Infinitival VP Complements in French".

This report finishes with general conclusions summing up the preceding parts and with a few suggestions concerning the near future of some of the tasks of the ACORD Project.

See also

Gabriel G. Bès, « Cliticos en frances y modelos linguisticos », *Revista Argentina de Lingüistica*, 2(2), 1986, p. 246-265.

Karine Baschung, Gabriel G. Bès, Annick Corluy, Thierry Guillotin, "Auxiliaries and Clitics in French UCG Grammar", *Conference of the European Chapter of the Association for Computational Linguistics*, 1987, p. 173-178.

Gabriel G. Bès, « Clitiques et constructions topicalisées dans une grammaire GPSG du français », *Lexique 6*, Presses universitaires de Lille, 1988.

ESPRIT PROJECT 393 ACORD

Construction and Interrogation of Knowledge Bases
using Natural Language Text and Graphics

FEASIBILITY OF A GPSG FRENCH GRAMMAR

by Gabriel G. Bès
and Karine Baschung

to include in T2.2
Investigation to Adapt GPSG Parser
to French Grammar

November 30, 1985

Revised version with further comments
February 8, 1986

CL II
Université de Clermont II
34, Avenue Carnot
F-63037 Clermont-Fd Cedex

CONTENTS

INTRODUCTION

The goal of this report is the feasibility of specifying a French grammar within the formalism proposed in GPSG.

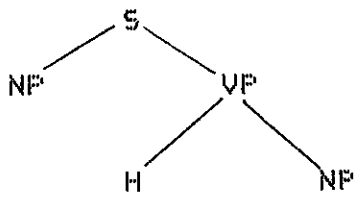
The linguistic specification of this formalism is revised in "A review of GPSG". As we feel that one of the less dangerous ways to evaluate a linguistic model is to try to specify particular grammars, a variety of French constructions were studied and specified with some details (though not in an exhaustive or definite manner) along the lines of the GPSG linguistic formalism; the condensed results are presented in "A GPSG French Grammar" and in "Some Remarks on Infinitival VP Complements in French".

This report finishes with general conclusions resuming the preceding parts and with a few suggestions concerning the near future of some of the tasks of the ACORD project.

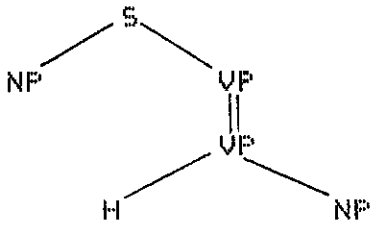
Notational Conventions

GPSG = Gerald Gazdar, Ewan Klein, Geoffrey Pullum and Ivan Sag, GENERALIZED PHRASE STRUCTURE GRAMMAR; Cambridge, Mass.; Harvard University Press, 1985.

Trees are sometimes dissociated in local trees. For example, the tree :



is noted



In some particular discussions, inherited features are noted : /NP and instantiated feature are noted : /NP.

A	:	admitted by the grammar	
¬A	:	not admitted by the grammar	
*	:	deviant (not grammatical)	Observational values
G	:	grammatical	

A REVIEW OF GPSG

In the following, GPSG will not be presented (see the deliverable on task 1.1). Instead, we condensed succinctly a number of points which offer some difficulties. Some of them are minor points, others seem somewhat more delicate to handle.

1. Linear order and ECPO. LP statements.

Trying to grasp all implications of ECPO from the inductive presentation of GPSG p. 47-49 seems difficult : notes 8 and 9 in chapter 3 and note 26 in chapter 5 are not transparent. The conceptual issue here is the determination of the strength of ECPO.

Ordered pairs are introduced for assigning values to variables in the syntax of coordination (GPSG p. 171 and note 6 in chapter 8). What is the relation between this possibility of incorporating order constraints in the grammar and ECPO?

It seems that there are linear order problems with the trees licensed by the following rules :

NP -> NP[+POSS], H[1]

AP -> (A2[+ADV]), H[1]

VP -> H[5], NP, NP (cf. Ditransitive constructions)

Cf. also SAI Metarule.

2. Ditransitive constructions

2.1 H[5] yields the following results :

(1) G A Lee handed Kim a book

(2) * A Lee handed a book Kim

(3) G A Peter sent a book to John

(4) G A Peter sent Mary to John

(5) G A Peter sent John a book

(6) * A Peter sent John Mary

Assuming the argumentation of GPSG p. 30-33, it is difficult to discriminate (2) and (6) from the other cases by

selectional features of a semantic or pragmatic nature. By application of the Passive Metarule, we can also have :

(7) * A A book was handed Kim

(8) * A John was sent Mary

2.2 Suggestion

The incorporation of a syntactic feature HUMAN {+,-} seems unavoidable; it is also required for prenominal possessive constructions. It may also be that what happens with (1) to (8) is quite close to clitic order in French (and other Romance languages). So we suggest the following LP statements :

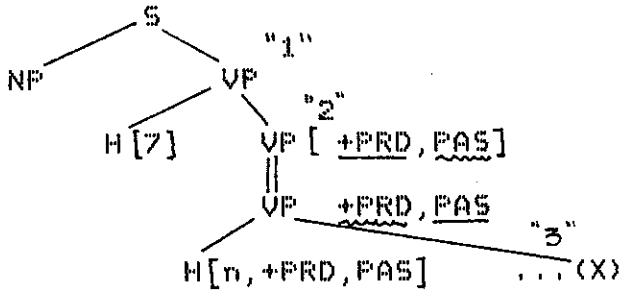
NP [+HUM] < NP [-HUM]

NP [+HUM] < NP [+HUM]

3. Passive Metarule

The typical tree-structure for passive forms is :

(1)



where "3" is the output of the passive metarule.

3.1 Problems about the local tree "3" of (1)

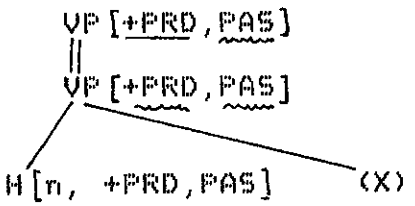
Since [+PRD] is a HEAD feature, we wonder if FCR 14 can force PAS onto the lexical head, which, if it is the case, will become exempt from its default value (FSD 7). Then, we can have PAS on VPs without being inherited from an output of the passive metarule :

* A Kim is proved the theorem

* A Kim is put the book on the table

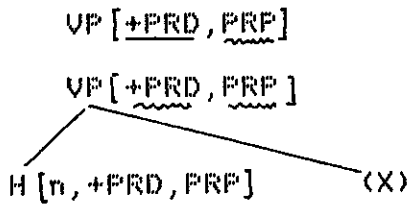
On the other hand, a tree like (2) :

(2)



does not, strictly speaking, satisfy the comments on definition 13 (ii) GPSG p. 103, because there exist as candidate projections trees like (3) :

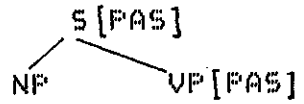
(3)



3.2 Problems about the local tree "1" of (1)

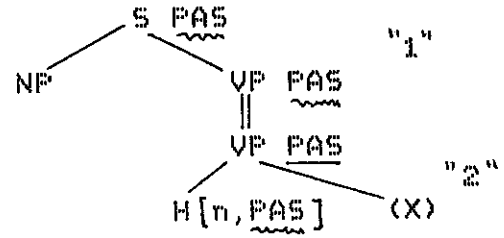
GPSG admits (see note 2 p. 73) :

(4)



"We assume that this is a wellformed S[VFORM PAS], but that the latter category plays no role in the grammar of English (i.e. no rule of English that introduces S either invokes it or permits it, and thus it cannot be introduced)". It is the case that no rule introducing S invokes PAS; however we cannot see anything preventing it from being freely instantiated. It seems that the tree (5) is admitted, where "2" is the projection of any output of the passive metarule :

(5)



If so, we have :

* A The theorem proved (by Kim)

* A The cake eaten (by Kim)

More generally, we can not see anything in GPSG forcing [FIN], so we have :

* A John love Mary

* A John to love Mary

3.2.1 Suggestion

Substitute FSD 7' for FSD 7 and add FSD n+1 and FSD n+2 :

FSD 7' : $\sim$ [PAS]

FSD n+1 : $\sim$ [BSE]

FSD n+2 : $\sim$ [INF]

3.3 Problems with some of the outputs of the Passive Metarule

The ID rules H[7] and H[44] satisfy the conditions on the input of the Passive Metarule, but the respective outputs are inadequate :

(a) VP [PAS] -> H[7], (PP[by])
* John is been (by Lee)

(b) VP [PAS] -> H[44, +it], S[+R], (PP[by])
*[It is] been who came (by Lee)

(c) VP[PAS] -> H[44, +it], S[FIN]/X2, (PP[by])
*[It is] been Kim loved (by Lee)

3.4 Passivization of an S

While the Lexical Rule for Passive Forms (GPSG p. 219) does mention the passivization of an S, the passive metarule does not and cannot, strictly speaking, generate strings like

That Lee is happy was said (by Lou)

where <u>was</u> must agree with an S.

3.4.1 Suggestion

Another Pas-Met and another ID rule for <u>be</u> can be worked along the lines of the following Pas-Met 2 and ID rule respectively :

Pas-Met 2 VP -> W, S[FIN]
$\Downarrow$
VP [PAS, AGR S] -> W

VP [+AUX] -> H[7 bis], VP[+PRD, AGR S]

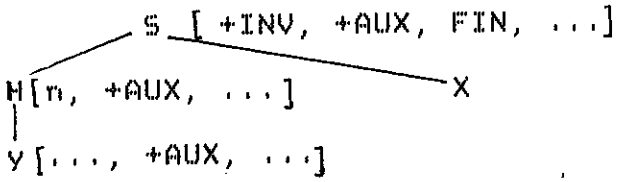
thus, we shall have :

That Lou was there yesterday is { true / interesting / believed by Kim }

4. SAI Metarule

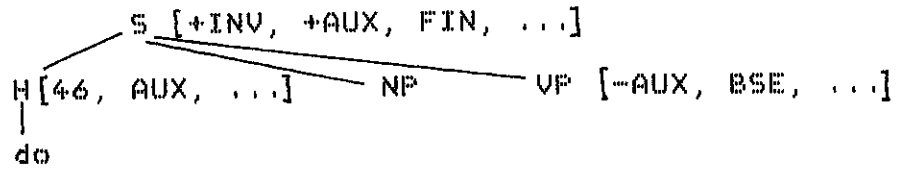
SAI and FCR 1 specify local trees as

(1)

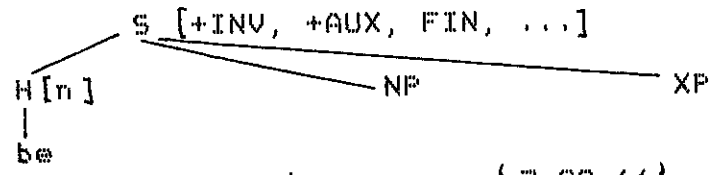


In GPSG only the ID rules H[7] and H [46] have the licensing feature [+AUX] in the mother, but we cannot see anything which forbids H[22] and H[44] becoming inputs to SAI : in any case, an "inversion" must apply to the structures associated with these two rules; moreover, there is a lexical entry be motivated by H[7] with [+AUX] and there seems to be no reason to mark the be of H[22] and H[44] as [-AUX] . So we assumed that SAI can apply to all and only H[7], H[22], H[44], and H[46], all other verbal lexical entries being marked [-AUX]. (Only VP [+it] -> H[44] , NP, S [+R] is considered here; questions related to VP [+it] -> H[44], X^2, S [+FIN] /X^2 were not treated). So we have the two schematic trees :

(2)



(3)



where n = {7,22,44}

4.1 Agreement problems in (2)

In (2) there must be an agreement relation between NP and, both H[46] and VP [-AUX]:

* Do John try ...

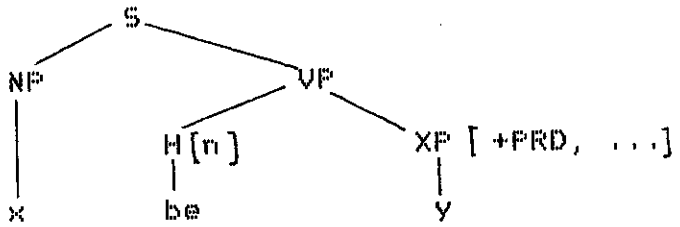
* Does it try ...

G Does it seem that ...

4.2 Agreement problems in the affirmative structures with be

The affirmative structures with be are associated with the schematic tree :

(4)



when XP = NP, there are problems with H[7] :

* A Mary is the brother of

* A John is the men ...

and with H[44] :

* A It is the men who

H[22] incorporates a rather ad-hoc solution (rule schemata) for number agreement.

4.3 Agreement and linear problems in the interrogative structures with be

Interrogative structures are associated with the schematic tree (3). They seem to cumulate problems related to affirmative structures (cf. § 4.2) and problems emerging by SAI application (cf. § 4.1) :

* Are John looking at himself

* Is John looking at herself

* Is John the sister of ...

* Is it the men who ...

But there are also some specific problems concerning linearity in interrogative structures with be. With H[7], XP may be NP, and must be NP with H[22] and H[44]. Thus, we have two NP sister categories in (3) with strict order relations :

G Is there a lion in the park

* Is a lion in the park there

G Is John the boy who proved the theorem

* Is the boy who proved the theorem John

In general, we cannot see in the present GPSG formulation of CAP and of LP statements the descriptive solutions for the above problems. It is probably not too difficult to provide a solution to some of them (LP statements). FCR 12 is an obstacle for handling agreement questions.

5. Specification of lexical entries

In general, it is difficult to formulate exhaustive lexical entries from what has been said about them in GPSG. This issue is related to the actual manipulation of FCR and FSD. We have some trouble with the differences between :

(a) nothing at all is noted about a feature and/or its value

(b) $\sim$ FEATURE X

(c) $\sim$ value x of FEATURE X

For concreteness. In the present formulations, AUX is a HEAD feature, but nothing prevents it from being present in NPs. What is the general policy here ?:

- Must all Ns be marked $\sim$[AUX] in lexical entries? Does this imply that all categories [+N, -V] must be written $\sim$[AUX] in local trees? Or will the convention be that nothing at all is noted in the trees and in lexical entries?

- Is a FCR required in these cases (say : [AUX]$\supset$ [+V, -N])?

This problem is related to our understanding of CAP. As it holds at present, we think that it is necessary to incorporate some counter intuitive lexical entries, as :

[INF, PER 3, -PLU] : to

[INF, PER 3, +PLU] : to

[INF, PER 2, -PLU] : to etc.

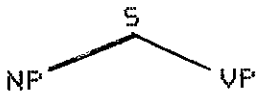
Same remarks for BSE.

Finally, is there some general policy about the choice of levels where it is possible to determine local tree admissibility ? IN GPSG [+AUX] is incorporated in the mother categories of ID H[7] and ID H[46] (dg), but it is absent from IDs H[22], H[44]. We think that if a FSD : ∼[AUX] is incorporated, and if [+AUX]becomes a licensing feature of mother categories of all the ID rules H[7], H[22], H[44] and H[46], the admissibility of trees can be determined without inspection of the leaves of terminated local trees; in this case, it does not seem necessary to mark all lexical entries other than H[7], [22], [44] and [46] as [-AUX]. They will be marked ∼[AUX] or nothing at all depending on notational conventions. In the present formulations, the admissibility is determined at the level of terminated local trees and lexical entries must be marked [-AUX] or [+AUX] (cf. also SAI Metarule).

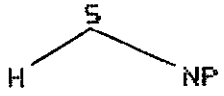
6. CAP

6.1 As it stands, GPSG's treatment of verb agreement seems to allow for the possibility of object agreement. We cannot distinguish between (1) and (2) by virtue of the definitions of Control and CAP, or via FCRs 12, 13 and FSDs 4, 10 :

(1)



(2)



Thus we may have strings like :

(3) A * I wants the book

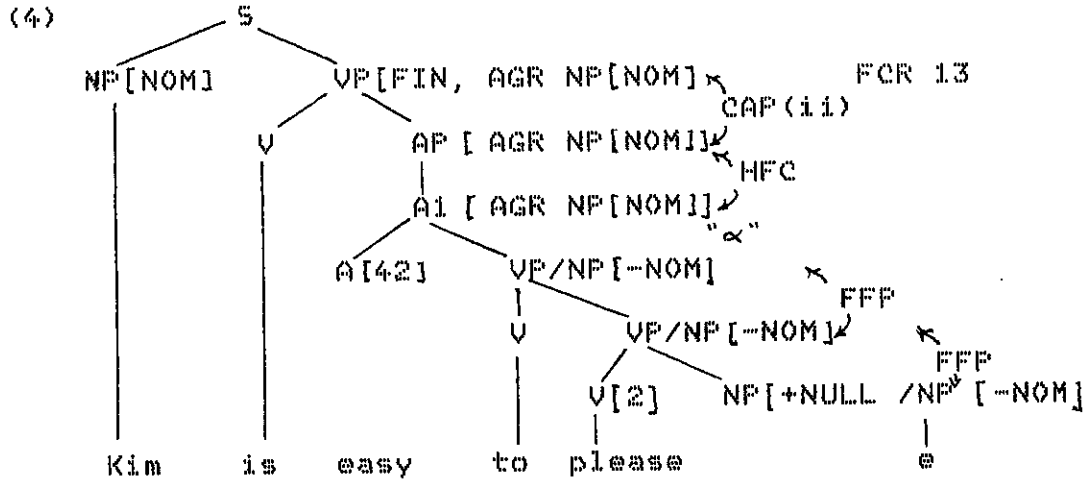
6.2 According to FCR 12 (GPSG (23) p. 88), [+V, +N] categories cannot bear the feature AGR in English. However, the discussion about missing object constructions invokes AGR on APs; we find p. 92 :

(34) A1 [AGR NP[*]]

A[42]

VP[SLASH NP[*]]

Strictly speaking, (34) is not admitted because A1 [AGR NP[*]] is not a legal extension. If we modify FCR 12 so as to admit AGR on APs (an issue that we must indeed assume in French and, we think, also in English for AP[+PRD], see SAI Metarule), a conflict arises between the values of control features in the local tree"α"of (4) :



In "α", CAP(ii) is not satisfied, thus the tree (4) cannot be admitted, which is incorrect.

7. Unbounded dependencies

The following observations concern only chapter 7 of GPSG. The consequences of UDC semantics have not been evaluated.

7.1 Problems with the following typical examples :

(1) * A Which Sandy reads

(2) * A The book of which Sandy reads

(3) * A { The house of / A discussion with } e Sandy likes e

(4) * A Sandy, we want e to succeed e

(5) * A Peter a discussion with e Lee likes e

(6) * A The books which the criticism of e was prejudicial to the commercialisation of e

(7) * A Kim's, I read a book of e (= GPSG (24 b) p. 148)

(8) G -A How difficult do you think this problem will be for the students to solve? (= GPSG (42 b) p. 152) (We do not know how to dislocate the AP constituent)

(9) G -A The man Fido chased returned (= GPSG (74 d) p. 162)

(10) G -A John is reading which books?

7.2 The status of <u>that</u> is not clear : complementizer, relative pronoun or both?

7.3 Problems with many outputs of STM 1

Outputs of ID rules (X^2 = NP)

(10) (a) H[5] G A A book, Lee gave Kim

(b) ?* A Kim, Lee gave a book

(11) H[20] ?* A Lou, that Robin came bothered

Outputs of outputs of Passive Metarule

(12) (a) H[5] G A A book, Kim was given by Lee

(b) G A By Lee, Kim was given a book

(c) * A Kim, a book was given by Lee

(d) * A By Lee, a book was given Kim

Outputs of outputs of SAI Metarule

(13) H[46] * A John, does e eat

(14) H[7] * A John, is e a man

(15) H[22] * A There, was e a lion in the park

(16) H[44] * A It is e Kim who relies on Sandy

(17) H[44] * A It is e on Kim that Sandy relies

Outputs from ID rules ($X^2 = V^2$)

(18) H[12] * A Go home, John wants to e

(19) H[8] G A That it was a good idea, John persuaded Lee

(20) H[48] * A And bought a book, John went

(21) H[14] ?* A For us to stand alone, John prefers

(Non exhaustive list; our grammatical judgements are not very clear).

Recapitulations :

(1) STM 1 seems too permissive

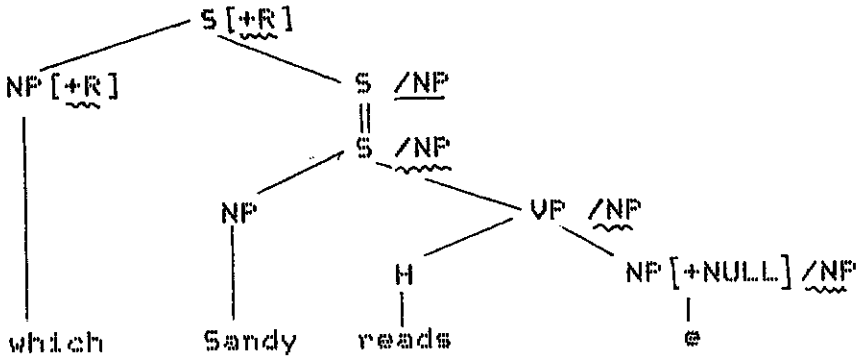
(2) The possibility of having two e seems adequate for parasitic gaps but is is also too permissive

(3) FCR 22 forbids echo questions

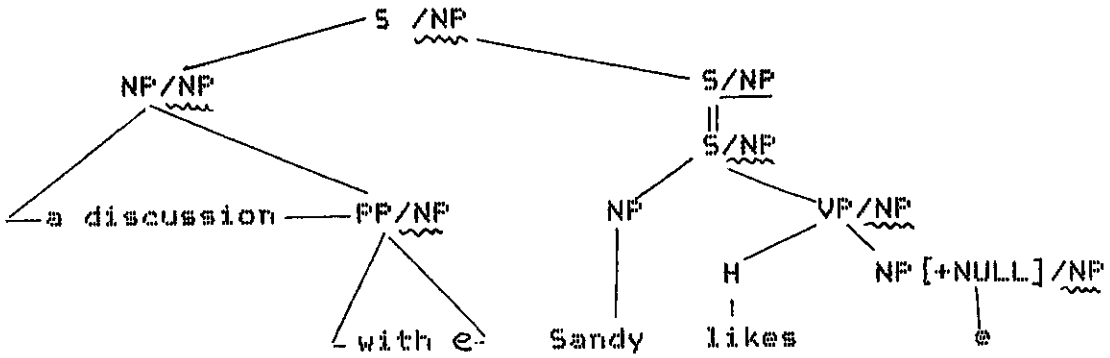
(4) [+Q] can be instantiated in a "root S" but this is not true for [+R]. But, we have difficulties in formally specifying in GPSG terms the notion "root S".

Some characteristic trees :

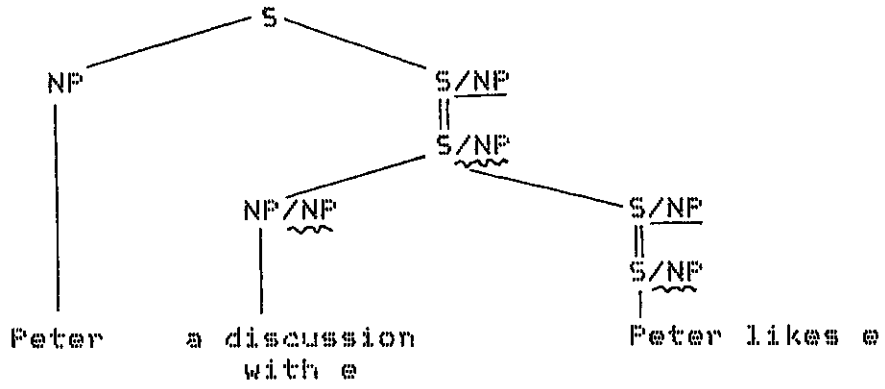
Associated with (1) :



Associated with (3) :



Associated with (5) :



8. Problems with the application of (22) Argument Order (GPSG p. 214)

If we understand (22), 'direct' object', by convention, equals '2-argument' and a meaning postulate is required for relating a 2-argument of give' with a 3-argument of give".

In constructions with free order constituents all of type NP, it seems that the application of (22) can introduce arbitrary decisions :

(1) I talk with Paul about Suzanne

In French (and Spanish) we have systematically :

(2) Il a donné un livre à Paul

(3) Il a donné à Paul un livre

and there are many verbs admitting both a direct and indirect object :

(4) toucher à ça

(5) toucher ça

In Spanish, we have

(6) Empujó à Pedro (He pushed Peter)

(7) Envió a Pedro (He sent Peter)

(8) Envió Pedro a María (He sent Peter to Mary)

(9) *Envió a Pedro a María

8.1 Suggestion

It seems that for questions relating to order arguments, TYP [NP ∼ PFORM] (the standard direct object) behaves differently than TYP [NP PFORMα]. It may be that the following principles are workable :

(I) If there is only one category of TYP NP (whether PFORM or PFORMα), it is the 2-argument

(II) If there is only one category of TYP[NP ∼ PFORM], it is the 2-argument

(III) If there is more than one category of TYP [NP ∼PFORM] , the first one is the 2-argument

(IV) If there is more than one category of TYP [NP PFORMα], the argument order is arbitrary

(I) can handle (4), (5), (6), (7).

(II) can handle GPSG (23) p. 214, (2), (3), (8).

(III) is a narrower reformulation of GPSG (22); it can handle (24).

(IV) can handle (1).

9. Metarules

Contrary to the specifications of the ProGram Manual (Evans & Gazdar 1984, p. 30; see also the report on Task 1.1 Evaluation of Different Parsers and Syntactic Theories p. 33), we cannot see anything in the GPSG specifications of the linguistic descriptive constraints themselves (GPSG p. 57-67) which forbids the application of a metarule to the rules which are outputs of this same metarule (but see GPSG p. 72). In some cases, this possibility seems empirically adequate :

(1) With whom and about what did Peter talk?

(2) Avec qui et à propos de quoi Pierre a-t-il discuté?

But if so, some other cases become problematic; for example, the Passive Metarule can apply twice to H[5], thus we may have as an output :

(3) VP[PAS] -> H[5], (PP[by]), (PP[by]).

A GPSG FRENCH GRAMMAR

1. Observational domain*

Interrogative inversion and clitic placement are the French constructions which the grammar in § 2 attempts to account for. The choice of the former one was obviously dictated by the ACORD tasks of Ld M and CL II. The latter is more specifically dictated by this report on the feasibility of specifying a French GPSG grammar. Because clitic placements seem at the outset to be particular constructions of Romance languages, or, at least, French constructions absent from English grammar, it appears that if it is possible to account for them within a GPSG formalism, a positive answer for the feasibility question of specifying a French GPSG grammar will be supported.

Interrogative inversion, clitic placement, left dislocation and/or topicalization seem to be interrelated constructions :

(1) Aime-t-il ce livre?
Does he like this book?

(2)* * Aime-Pierre ce livre?
Does Peter like this book?

(3) Pierre, aime-t-il ce livre?
Peter, does he like this book?

(4) Ce livre, Pierre l'aime
This book, Peter likes

In (3) and (4), pronouns corefer with an NP in a topicalized position; in the inverted position there must be a pronoun, otherwise the construction is deviant (cf (2)).

Other interrogative French constructions present left movement, but with different consequences than (3) on coreferential pronouns :

(5) (a) Quel livre Pierre aime?
Which book does Peter like?

(b) *Quel livre Jacques l'aime?

Observations on (1) to (5) finally determined the main constructions aimed at : (a) (a subclass, see below) of interrogative inversion; (b) clitics, (c) topicalized constructions and wh-questions in root sentences; (d) the classic passive was added to these constructions.

* § 1 benefits from the collaboration of N. Rousseau-Payen.

The main observations which concern these constructions and which the grammar of § 2 can tackle are systematized in § 1.1 to 1.5.

1.1 Pronoun inversion

In French there are several kinds of "inversion" (this term and others like movement are used throughout this report without a technical sense) other than (1) and (3). The report on Task 2.3 will present a detailed overview of interrogative French constructions ; among them :

(6) (a) Le livre que regarde Pierre ...
The book which Peter looks at

(b) Quel livre souhaite acheter Pierre?
Which book does Peter want to buy?

Inversions as in (6 a) or (6 b) are not accounted for by the grammar of § 2, but it is hoped that a more complete formulation of LF statements can overcome these problems and more generally the somewhat intricate questions of French order (see Conclusions below).

Pronoun inversion - the subclass of French inversions aimed at here - has the effect of placing a pronoun (1st, 2nd or 3rd) to the right of the tensed verb form, a main verb or an auxiliary, active or passive :

(7) Connais-tu le livre?
Do you know the book?

(8) A-t-il aimé le livre?
Did he like the book?

(9) Pierre, a-t-il aimé le livre?
Peter, did he like the book?

(10) (Pierre), a-t-il été aimé?
(Peter) was he loved?

The optative NP in a topicalized position must agree in gender, number and person with a pronoun in the main sentence; this pronoun may be the inverted pronoun. Exceptions with PER 3 and PER 2 are only apparent (cf. (12)): a PER 3 NP to the left of the main S functions in this case as a sort of vocative which can optionally be preceded or followed by a PER 2 NP in a topicalized position :

(11) * Pierre, a-t-elle aimé le livre?

(12) Pierre, connais-tu le livre?
Peter, do you know the book?

(13) Toi, connais-tu le livre?

You, do you know the book?

(14) Toi, Pierre, connais-tu le livre?
You, Peter, do you know the book?

(15) Pierre, toi, connais-tu le livre?

Constructions as in (14) and (15) are not accounted for by the grammar of § 2.

Interrogative pronouns are excluded from interrogative inversion :

(16) * Pierre connaît-qui le livre?

1.2 Clitics

(a) Clitics as PRO-forms. Some clitics can corefer with an NP in a topicalized position ((17) and (18)), but this NP can be absent; others must resume a construction in this same position ((19) and (20)) but, in this case, with a different semantic relation than that in (17) or (18). According to the relation between the clitic element with the constituent in the topicalized position and/or in the VP (see below), clitics are not always stricly speaking PRO-NOUNS : they can also be PRO-AP (19), PRO-PP (21), PRO-S (22-a) and PRO-VP (22-b) :

(17) (Pierre) Marie l'aime
(Peter) Mary loves him

(18) (Toi) Marie t'aime
(You) Mary loves you.

(19) Intelligent, Pierre l'est
Intelligent, Peter is

(20) * Pierre l'est

(21) A la poste, Pierre y va tous les jours
To the post-office, Peter goes daily

(22) (a) Qu'il est déjà tard, Pierre le sait
That it is late, Peter knows

(b) Connaître la vérité, Pierre le veut
To know the truth, Peter wants it

(b) Clitic placements. Clitics are particular forms (see (d) below) that can be placed in front of the VP, to the left of a tensed verb form, a main verb or an auxiliary, active or passive (23), or to the left of an infinitive form or of a non adjectival participle (24); more than one clitic can be

placed in this position (25).

(23) (a) Marie le regarde
Mary watches {him, it}

(b) Marie l'a regardé
Mary watched {him, her, it}

(c) Il en a parlé
He talked about it

(d) Le livre lui a été donné
The book was given to {him, her}

(24) (a) Marie dit le connaître
Mary says that she knows {that, him}

(b) Le lui disant souvent ...
Saying it to him often ...

(25) (a) Pierre le lui a donné
Peter gave it to him

(b) Pierre lui en a parlé
Peter talked to him about it

Grammar of § 2 does not account for clitics as in (24 b).

(c) Clitics and VP constituents. There are clear relations between the VP constituents placed to the right of the verbal forms and the clitic forms admitted by the same verbs; moreover, one or more complements can be optionally cliticized ((26), (27) and (28)), but it is not true that all VP constituents can be cliticized (compare (29) to (28)) :

(26) (a) Pierre a présenté {un ami, un cadeau} à Marie
Peter presented {a friend, a gift} to Mary

(b) Pierre le lui a présenté
Peter presented {it, him} to {him, her}

(c) Pierre l'a présenté à Marie
Peter presented {it, him, her} to Mary

(d) Pierre lui a présenté {un ami, un cadeau}
Peter presented {a friend, a gift}
to {him, her}

(e) * Pierre le y a présenté

(27) (a) Pierre raffole de cela
Peter enjoys that

(b) Pierre en raffole

(c) * Pierre {le, y} raffole

(28) (a) Pierre pense à la promenade
Peter thinks of the walk

(b) Pierre y pense

(29) (a) Pierre pense à lui
Peter thinks of him

(b) * Pierre lui pense

(d) Order relations among clitics. Clitic forms are : {{me, nous, te, vous, se}A, {le, l', la, les, lui, leur, y, en}B}. (The partition will be obtained by the feature REF, see p. 35). There are strict order relations between clitics, which can be synthetized in (30) to (33); they are respectively illustrated in (34) to (37).

(30) A clitic from class A precedes a clitic from class B.

(31) A clitic from class A cannot be followed by lui

(32) Le precedes lui, y

(33) Lui, y precedes en

(34) (a) Pierre t'en a parlé
Peter talked to you about that

(b) *Pierre en t'a parlé

(c) Pierre me le porte
Peter brings it to me

(d) * Pierre le me porte

(35) (a) Pierre t'a présenté à lui
Peter presented you to him

(b) * Pierre te lui a présenté

(36) (a) Pierre le lui a donné

(b) * Pierre lui le a donné

(c) Pierre l'y a envoyé
Peter sent {it, him} there

(d) * Pierre y l'a envoyé

(37) (a) Pierre lui en a parlé
Peter talked to him about that

(b) * Pierre en lui a parlé

Clitic order is quite different from other constituents order. In a topicalized position, the NP order seems quite free (38), as it is the order of "ordinary" VP constituents (39), in contrast with the rigid clitic order presented above; moreover, in imperative constuctions some (but not all) "clitic" forms appear in a non clitic position (i.e., to the right of a verb form), but, in standard French, with a rigid order different from the ordinary constituents one and different also from the rigid order of the corresponding forms in clitic position (40); these constructions present moreover a neat substandard variation whose order is very close to the standard rigid order of the clitic position, (the grammar in § 2 does not account for imperative order) :

(38) (a) A Marie, le journal Pierre le lui porte tous les jours
To Mary, the journal Peter takes to her daily

(b) Le journal, à Marie Pierre le lui porte tous les jours

(39) (a) Pierre porte à Marie le journal tous les jours

(b) Pierre porte le journal à Marie tous les jours

(40) (a) Porte le moi
Bring it to me

(b) * Porte moi le (standard)

(c) G Porte moi le (sub-standard)

(e) Morphological variations. Personal pronouns present morphological variation according to the occupied position : topicalized, subject, clitic or VP constituent :

(41) (a) Toi tu penses à moi
You you think of me

(b) * Tu toi penses à me

(c) Moi je te le dis
I say it to you

(d) * Me, moi tu le dis

1.3 Clitics and inverted pronouns

Clitics and pronoun inversion can be combined in the same sentence, and characteristic features of each construction all seem to be preserved :

(43) (a) La connais-tu?
Do you know her?

(b) Toi, la connais-tu?
You, do you know her?

(c) * Pierre, la connaît qui?

(d) Pierre, le lui a-t-il donné?
Peter, did he give it to him?

(e) Le livre, lui a-t-il été donné?
Was the book given to him?

1.4 Topicalized constructions and wh-questions

A variety of constructions (see above (17) to (22)) can appear in topicalized position. Relations between the topicalized constructions and S constituents (including subjects) can be synthetized as in (44) and (45); they are illustrated in (46) and (47), respectively :

(44) A non-interrogative (=¬Q) topicalized construction must be resumed by a phonic (i.e. overt) PRO-form, clitic or non clitic, in the main S (the coreference relation between NP and PP with PRO-NP and PRO-PP respectively being a particular case of the more general "resuming relation").

(45) An interrogative (= + Q) topicalized construction (wh-questions in root sentences) can not be resumed by an overt PRO-form.

(46) (a) Pierre, Marie le connaît
Peter, Mary knows {him, it}

(b) Pierre, il la connaît
Peter, he knows {her, it}

(c) (A) Pierre, Marie lui a donné un livre
(To) Peter, Mary gave a book to him

(d) * Pierre, Marie connaît ce livre

(e) A la poste, Pierre y va tous les jours
To the post-office, Peter goes daily

(f) * A la poste, Pierre va tous les jours

(g) *? Pierre, Marie connaît
Peter, Mary knows

(h) Que tu connaisses la raison, Pierre le sait
That you know the reason, Peter knows it

(i) * Que tu connaisses la raison, Pierre sait

(47) (a) Quel livre Pierre connaît?
What book does Peter know?

(b) * Quel livre Pierre le connaît?

(c) A qui Pierre a donné un livre?
To whom did Peter give a book?

(d) * A qui Pierre lui a donné un livre?

(e) Qui parle?
Who is speaking?

(f) * Qui il parle?

There are dialectal variations on judgements about some constructions related to (44), see (46 g). The grammar of § 2 accounts for the conservative view of standard French expressed in (44). Moreover, the grammar of § 2 accounts only for an overt prepositional dative construction in a topicalized position (see (46 c)).

1.5 Topicalization, wh-questions and clitics

It has been pointed out (see paragraph 1.3) that topicalized ∼Q constructions, pronoun inversion and clitics can appear in one sentence. The same is true for topicalized +Q constructions (48); +Q NP subjects can coexist with clitics (49) :

(48) (a) Quel livre lui-a-t-il donné?
Which book did he give to {him, her}?

(b) A qui a-t-elle été recommandée?
To whom was she recommended?

(c) Avec quoi l'a-t-il transporté?
With what did he transport it?

(49) (a) Qui l'a transporté?
Who transported it?

(b) La table, qui l'a transportée?
The table, who transported it?

In refined standard French, a topicalized +Q

construction requires pronoun inversion in the main sentence ((50 a) and not (50 b)), and echo questions (51) are not accepted. But forms such as (50 b) and (51) are very common in a more colloquial standard French and the grammar in § 2 can in principle account for them :

(50) (a) Quel livre a-t-il acheté?

(b) Quel livre il a acheté?
Which book did he buy?

(51) Il a acheté quel livre?
He bought which book?

2. French grammar

The following GSPG French grammar attempts to account for the observational domain outlined in § 1. It is by no means neither a definite nor a complete grammar. In addition to the limitations mentioned in § 1 another must be added : PER 3 clitic form se presents particular problems and is outside the scope of the solutions proposed below. Other limitations will be mentioned in paragraph 3.

The proposed grammar was designed as close as possible to the English grammar of GPSG. Comments on the results will be presented in § 3.

2.1 Features and feature values

feature :	value range :
COMP	{que, si}
NULL	{+,-}
WHMOR	{Q,...}

HEAD features

feature :	value range
AGR	CAT
BAR	{0, 1, 2}
CAS	{NOM, ACC, OBL(ique)}
CLIT(ic)	{+, -}
DET(erminer)	{DEF(inite), IND(efinite), INT(errogative)}
GEND(er)	{MASC, FEM}
HUM(an)	{+, -}
INV(ersion)	{+, -}
N	{+, -}
PAS	{+, -}
PER	{1, 2, 3}
PFORM	{à, de, en, avec}

PROF(orm)	{PRON(oun), PROVP, PROAP, PROS}
PLU	{+, -}
REF(erential)	{+, -}
SLASH	CAT
SUBCAT	{1, ..., n} ∪ {gue, si, ...}
SUBJ	{+, -}
TENSE	{PAST, PRES, ...}
V	{+, -}
VFORM	{FIN, INF}

FOOT features

feature :	value range
SLASH	CAT
WH	CAT

2.2 Feature co-occurrence restrictions

FCR 1 : [PRON, PER 1] ⊃ [+REF, +HUM]

FCR 2 : [PRON, PER 2] ⊃ [+REF, +HUM]

FCR 3 : [+CLIT] ⊃ [PROF]

FCR 4 : [PROF] & [NP] ⊃ [PRON]

FCR 5 : [PROF] & [VP] ⊃ [+CLIT, PROVP]

FCR 6 : [PROF] & [S] ⊃ [+CLIT, PROS]

FCR 7 : [PROF] & [AP] ⊃ [+CLIT, PROAP]

FCR 8 : [+CLIT] ⊃ [REF] & ~[CAS]

FCR 9 : [PRON] ⊃ [PER] & [+N, -V]

FCR 10 : [PFORM] ⊃ ([NP] ∨ [PP])

FCR 11 : [INT] ⊃ [+Q]

FCR 12 : [+INV] ⊃ [FIN] & ~[COMP]

FCR 13 : [S/NP[PRON, +REF]] ⊃ [S/NP[PRON, +REF, OBL]]

FCR 14 : [S/XP[+Q]] ⊃ [+INV](depending on dialects)

FCR 15 : [VFORM] ⊃ [+V, -N]

FCR 16 : [BAR 0] = [N] & [SUBCAT] & [V]

FCR 17 : [BAR 1] = ~[SUBCAT]

FCR 18 : [BAR 2] = ~[SUBCAT]

FCR 19 : [+INV, BAR 2] ⊃ [+SUBJ]

FCR 20 : [+SUBJ] ⊃ [+V, -N, BAR 2]

FCR 21 : [AGR] ⊃ [-N, +V]

FCR 22 : [FIN, AGR NP] ⊃ [AGR NP[NOM]]

FCR 23 : [+NULL] ⊃ [SLASH]

FCR 24 : [+PAS] ⊃ [+V, -N]

2.3 Feature specification defaults

FSD 1 : [FIN]

FSD 2 : ~[NULL]

FSD 3 : ~[NOM]

FSD 4 : [PFORM] ⊃ [BAR 0]

FSD 5 : [BAR 0] ⊃ ~[PAS]

FSD 6 : [+N, -V, BAR 2] = [ACC]

FSD 7 : ~[REF]

FSD 8 : [-CLIT]

FSD 9 : [+N, -V] ⊃ ~[PFORM]

FSD 10 : ~[/[XP, ~Q]]

2.4 Immediate dominance rules

Non lexical ID rules

$S \rightarrow X^2, H\ [-SUBJ]$

$S\ [COMP\ \alpha] \rightarrow \{[SUBCAT\ \alpha]\},\ H\ [COMP\ NIL, -Q]$
where α is in {que, si}

$S \rightarrow X^2, H/X^2$

$AP \rightarrow H^1$

$PP \rightarrow P^1$

$NP\ [DET\ \alpha] \rightarrow \{[SUBCAT\ \alpha]\}, N^1$
where α is in {DEF, IND, INT}

Lexical ID rules other than VP :

NP [à, -CLIT] -> P [1], NP [~PFORM, OBL] (à)

NP [de, -CLIT] -> P [2], NP [~PFORM, OBL] (de)

NP [avec, -CLIT] -> P [3], NP [~PFORM, OBL] (avec)

P^1 -> H [4], NP [OBL] (sur, dedans, dans ..)

NP -> H [5] (Peter, Mary)

NP -> H [6] (je, tu, il ...)

N^1 [+HUM] -> H [7] (garcon, fille)

N^1 [-HUM] -> H [8] (cadeau, livre)

A^1 -> H [9] (intrépide)

VP Lexical ID rules :

VP -> H [10], NP, NP [à, +HUM] (présenter)
- une amie au garcon

VP -> H [11], NP, NP [à, -HUM] (présenter)
- une lettre au guichet

VP -> H [12], NP [+PL] (grouper)
- les enfants, *- l'enfant

VP -> H [13], NP, NP [avec] (grouper)
- la fille avec le garcon

VP -> H [14], NP [à, +HUM] (plaire)
- à la fille

VP -> H [15], NP [à, -HUM] (remédier)
- à la douleur

VP -> H [16], NP [à, +HUM] (appartenir)
- à Marie

VP -> H [17], NP [à, -HUM] (appartenir)
- à cet engin

VP -> H [18], NP [à, +HUM, -CLIT] (penser)
- à lui *- lui penser

VP -> H [19], NP [à, -HUM] (penser)
- à la promenade - y penser

VP -> H [20], NP, NP [à, -CLIT] (habituer)
- Pierre à Paul *- lui habituer Pierre

VP → H [21], NP (voir)
- le cadeau

VP → H [22], NP [de] (raffoler)
- de cela

VP → H [23], NP, NP [à, +HUM] (enlever)
- une douleur à quelqu'un

VP → H [24], NP, NP [de] (convaincre)
- Pierre de son erreur

VP → H [25], AP (être)
- intrépide

VP → H [26], NP, PP (mettre)
- un livre sur la table

VP → H [27], S [COMP que] (penser)
- que Pierre viendra

VP → H [28], S [COMP si] (ignorer)
- si Pierre viendra

2.5 Metarules

Passive Metarule

$$\begin{array}{lcl} VP & \rightarrow & W, NP \\ \Downarrow & & \\ VP[PAS] & \rightarrow & W, (NP[par]) \end{array}$$

Clitic Metarule

$$\begin{array}{lcl} V^2 & \rightarrow & W, X^2 \\ \Downarrow & & \\ V^2 & \rightarrow & W, X^2\ [+CLIT] \end{array}$$

Pronoun-Inversion Metarule

$$\begin{array}{lcl} VP & \rightarrow & W \\ \Downarrow & & \\ V^2[+INV] & \rightarrow & W \end{array}$$

Slash-Termination Metarule

$$\begin{array}{lcl} X & \rightarrow & W, X^2 \\ \Downarrow & & \\ X & \rightarrow & W, X^2\ [+NULL] \end{array}$$

2.6 LP statements

LP 1 : [+CLIT] < [SUBCATα, -CLIT]

LP 2 : [+CLIT, +REF] < [+CLIT, -REF]

LP 3 : [+CLIT, +HUM] < [+CLIT, +HUM]

LP 4 : [+CLIT, NP, $\sim$PFORM, -REF] < [+CLIT, NP, PFORMα, -REF]

LP 5 : [+CLIT, -REF, à] < [+CLIT, -REF, de]

LP 6 : [SUBCATα, -CLIT] < [$\sim$SUBCAT, -CLIT]
and it is also assumed (cf Baschung § 2)

LP 7 : [NP] < VP[$\sim$INF]

2.7 Examples for lexical entries

Verbal entries :

LE 1 : <présente, [[-N], [+V], [BAR 0], [SUBCAT 10], -INV, -PAS, FIN, PRES, AGR NP [PER 3, -PL, [GENDER]], $\sim$/]>

LE 2 : <présentent,[[-N], [+V], [BAR 0], [SUBCAT 10], -INV, -PAS, FIN, PRES, AGR NP [PER 3, +PL, [GENDER]], $\sim$/]>

LE 3 : <a présenté, [[-N], [+V], [BAR 0], [SUBCAT 10], -INV, -PAS, FIN, PAST, AGR NP [PER 3, -PL, [GENDER]] , $\sim$/]>

LE 4 : <a été présenté, [[-N], [+V], [BAR 0], [SUBCAT 10], -INV, +PAS, FIN, PAST, AGR NP [PER 3, -PL, [GENDER]], $\sim$/]>

LE 5 : <a-t-il été présenté, [[-N], [+V] , [BAR 0] , [SUBCAT 10], +INV, +PAS, FIN, PAST, AGR NP [PER 3 , -PL, MASC ,$\sim$/]>

LE 6 : <a-t-il été présenté, [[-N], [+V], [BAR 0], [SUBCAT 10], +INV, +PAS, FIN, PAST, AGR NP [PER 3 , -PL, MASC , /NP [PER 3, -PL, MASC, $\sim$Q] >

LE 7 : <présente-t-il, [[-N], [+V], [BAR 0], [SUBCAT 10], +INV, -PAS, FIN, PAST, AGR NP [PER 3, -PL, MASC ,$\sim$/]>

LE 8 : <présente-t-il, [[-N], [+V], [BAR 0], [SUBCAT 10], +INV, -PAS, FIN, PAST, AGR NP [PER 3, -PL, MASC], /NP [PER 3, -PL, MASC, $\sim$Q] >

LE 9 : ‹a-t-il présenté, [[-N], [+V], [BAR 0], [SUBCAT 10], +INV, -PAS, FIN, PAST, AGR NP [PER 3, -PL, MASC], ~/]›

LE 10 : ‹ a-t-il présenté, [[-N], [+V], [BAR 0], [SUBCAT 10], +INV, -PAS, FIN, PAST, AGR NP [PER 3, -PL, MASC], /NP [PER 3, -PL, MASC, ~Q]›

Pronoun entries :

LE 11 : ‹tu, [+N, -V , [BAR 0] , [SUBCAT 6], PRON, PER 2, -PL, ~ GEN, +HUM, -CLIT, +REF, ~ PFORM, ~ WH, NOM, ~/]›

LE 12 : ‹tu, [+N, -V , [BAR 0] , [SUBCAT 6], PRON, PER 2, -PL, ~GEN, +HUM, -CLIT, +REF, ~ PFORM, ~ WH, NOM, /NP[PER 2] ›

LE 13 : ‹toi, [+N, -V , [BAR 0] , [SUBCAT 6], PRON, PER 2, -PL, ~ GEN, +HUM, -CLIT, +REF, ~ PFORM, ~ WH, OBL, ~/]›

LE 14 : ‹te, [+N, -V , [BAR 0] , [SUBCAT 6], PRON, PER 2, -PL, ~ GEN, +HUM, +CLIT, +REF, ~ PFORM, ~ WH, ~CAS, ~/]›

LE 15 : ‹te, [+N, -V , [BAR 0] , [SUBCAT 6], PRON, PER 2, -PL, ~ GEN, +HUM, +CLIT, +REF, [PFORM à], ~ WH, ~CAS, ~/]›

LE 16 : ‹te, [+N, -V , [BAR 0] , [SUBCAT 6], PRON, PER 2, -PL, ~ GEN, +HUM, +CLIT, +REF, ~PFORM, ~ WH, ~CAS, /NP[PER 2] ›

LE 17 : ‹l(e), [+N, -V , [BAR 0] , [SUBCAT 6], PRON, PER 3, -PL, MASC, +HUM, +CLIT, -REF, ~ PFORM, ~ WH, ~CAS, ~/]›

LE 18 : ‹l(e), [+N, -V , [BAR 0] , [SUBCAT 6], PRON, PER 3, -PL, MASC, +HUM, +CLIT, -REF, ~ PFORM, ~ WH, ~CAS, /NP[PER 3, -PL, MASC, ~ Q] ›

LE 19 : ‹la, [+N, -V , [BAR 0] , [SUBCAT 6], PRON, PER 3, -PL, FEM, +HUM, +CLIT, -REF, ~ PFORM, ~ WH, ~CAS, ~/]›

LE 20 : ‹la, [+N, -V , [BAR 0] , [SUBCAT 6], PRON, PER 3, -PL, FEM, -HUM, +CLIT, -REF, ~ PFORM, ~ WH, ~CAS, ~/]›

LE 21 : ‹les, [+N, -V , [BAR 0] , [SUBCAT 6], PRON, PER 3, +PL, FEM, +HUM, +CLIT, -REF, ~ PFORM, ~ WH, ~CAS, ~/]›

LE 22 : ⟨les, [+N, -V , [BAR 0] , [SUBCAT 6], PRON, PER 3, +PL, MASC, +HUM, +CLIT, -REF, ~PFORM, ~WH, ~CAS, ~/]⟩

LE 23 : ⟨les, [+N, -V , [BAR 0] , [SUBCAT 6], PRON, PER 3, +PL, MASC, -HUM, +CLIT, -REF, ~PFORM, ~WH, ~CAS, ~/]⟩

LE 24 : ⟨lui, [+N, -V , [BAR 0] , [SUBCAT 6], PRON, PER 3, -PL, FEM, +HUM, +CLIT, -REF, PFORM à, ~WH, ~CAS, ~/]⟩

LE 25 : ⟨lui, [+N, -V , [BAR 0] , [SUBCAT 6], PRON, PER 3, -PL, MASC, +HUM, +CLIT, -REF, PFORM à, ~WH, ~CAS, ~/]⟩

LE 26 : ⟨y, [+N, -V , [BAR 0] , [SUBCAT 6], PRON, PER 3, -PL, MASC, -HUM, +CLIT, -REF, PFORM à, ~WH, ~CAS, ~/]⟩

LE 27 : ⟨y, [+N, -V , [BAR 0] , [SUBCAT 6], PRON, PER 3, +PL, MASC, -HUM, +CLIT, -REF, PFORM à, ~WH, ~CAS, ~/]⟩

LE 28 : ⟨y, [+N, -V , [BAR 0] , [SUBCAT 6], PRON, PER 3, +PL, FEM, -HUM, +CLIT, -REF, PFORM à, ~WH, ~CAS, ~/]⟩

LE 29 : ⟨y, [+N, -V , [BAR 0] , [SUBCAT 6], PRON, PER 3, -PL, FEM , -HUM, +CLIT, -REF, PFORM à, ~WH, ~CAS, ~/]⟩

Proform entries other than pronoun :

LE 30 : ⟨l(e), [... PROS, +CLIT, -REF, ~PFORM, ~WH, ~CAS, /S]

LE 31 : ⟨l(e), [... PROAP, +CLIT, -REF, ~PFORM, ~WH, ~CAS, /AP]

LE 32 : ⟨l(e), [... PROVP, +CLIT, -REF, ~PFORM, ~WH, ~CAS, /VP [INF]]

Interrogative entries :

LE 33 : ⟨qui, [+N, -V, [BAR 0], [SUBCAT 6], PER 3, +Q ..]⟩

LE 34 : ⟨quel, [...[SUBCAT INT], INT, +Q ..]⟩

LE 35 : ⟨e, [XP [+NULL]/XP, +Q] ⟩

3. Comments

The grammar of § 2, though not a complete or definite version is intended to cover the observational domain outlined in § 1. In what follows, the main points of § 1 are re-presented in relation to the proposed descriptive solutions

3.1 Pronoun inversion

Pronoun inversion is obtained without literally inserting a pronoun by, say, a metarule. This possibility has indeed been explored, but it leads to many inconveniences because, as it was pointed out in § 1.1, the pronoun must appear to the right of the tensed verbal form, which can be an auxiliary (avoir or être) or the main verb. The proposed solution takes advantage of GPSG formalism on morphology and of the fact that there is only a closed and very limited number of pronouns that can appear in this position; this is indeed a great difference with other French inversions (see § 1.1) and with the observational domain covered by the SAI Metarule of GPSG in English. In order to implement this solution on pronoun-inversion it is necessary to introduce lexical entries as[3]to[10]besides the more classical ones [1] and[2].

The distribution of [+INV] is severely limited. Thanks to ID rule S → X^2, H[-SUBJ] and FCR 19 it is impossible to have [+INV] in an ordinary subject plus predicate sentence. Because INV and PAS are not values of the same feature, an output of the Passive Metarule can be the input of the Pronoun-Inversion Metarule. All these seem to have empirically adequate consequences :

(1) (a) *Jean ignore si regarde-t-il le livre
(b) *Jean pense que regarde-t-il le livre

(2) A-t-il été aimé?

In the proposed grammar of § 2, [+INV]is introduced by a rather simple metarule, but it is hoped that in a simpler version even this metarule can be eliminated altogether : it may be that [INV] (as NEG(ative), a feature not incorporated to the commented grammar) is nothing more than a morphological verb feature with syntactic characteristics entirely determined by feature specifications (FCR and FSD).

Note that the main S to the right of a topicalized constituent may be [+INV]or [-INV], as required (on this point, see also § 3.5 below) :

(3) (a) Pierre, a-t-il lu le livre?
(b) Pierre, il a lu le livre

(4) (a) Intelligent, il l'est vraiment
(b) Intelligent, l'est-il vraiment?

(5) (a) Que Pierre connaisse le problème, elle le sait
(b) Que Pierre connaisse le problème, le sait-elle?

3.2 Clitics

The Clitic Metarule introduces [+CLIT] on BAR 2 constituents, and FCR 3 requires the resultant category to be a PROF(orm). The distribution of the different values of PROF is fixed by FCRs 4 to 7. Note that, as required, pronouns (i.e. [PROF PRON]) may or may not be [+CLIT]but categories with other values of PROF must be [+CLIT].

The Clitic Metarule and the LP statements 1 to 5 constitute the core of the solution on clitics proposed by the grammar of § 2; conjointly they are intended to account for clitic placements, clitic relations between clitic categories and VP non clitic constituents and order relations among clitics.

When [+CLIT]is inserted into a XP constituent by the Clitic Metarule, this category must obey LP 1 and it is thus placed to the left of the verbal head. A metarule is assumed to apply freely to the output obtained by this same metarule (for discussion, see A Review of GPSG § 9), so the number of clitic constituents is automatically determined; (6) is not admissible because there is no ID rule (7) :

(6) * Il en le revient

(7) VP --> H[n], NP, NP[de] (revenir)

The information placed by ID rules on VP constituents - information that, in any case, is necessary for subcategorizing verbal heads - is thus preserved. For example, for H[12],H[13],and H[14],the following results are obtained, respectively :

(8) (a) G A ... les grouper
(b) * -A ... le (la) grouper

(9) (a) G A ... le (la) grouper avec Pierre
(b) * -A ... le (la) avec Pierre grouper

(10) (a) G A ... lui plaît
(b) * -A ... y plaît

The result of (8 b) is obtained because H[12]requires a plural NP; the one of (9 b) because there is no lexical clitic entry with [...PFORM avec ...] and the one of (10 b), because the clitic category is ... [PFORM à, +HUM ...] and

not [... PFORM a, -HUM ...].The required lexical entries for (8 a), (9 a), and (10 a) are LEs[21] to [23], LEs[17] to [20], and LEs[24] or [25], respectively.

Besides the order relation between clitic and non clitic categories (handled by LP 1 and 6) there are strict order relations among clitic categories. These are handled by LP statements 2 to 5, which heavily rely on features HUM and REF.

The feature HUM is in any case needed for verb subcategorization and the distribution of forms such as lui et y. The feature REF(erential) seems to be an ad-hoc feature that enables the partition of the set of clitic forms, { me, nous, te, vous, se } being [+REF] and {le, l', la, les, lui, leur, y, en } being [-REF]. But it is hoped that REF is something more; "+REF(erential)" must be understood as something like : the referent is conjointly determined by the linguistic forms and the speech act (in the special case of se, there is obligatory coreference with the subject, but see the proviso at the beginning of paragraph 2). On the other hand "-REF(erential)" does not mean "without reference" but something like "with an open referent that must be different from that of the subject". FCR 8 requires that clitic categories must have a value for REF (and so LP2 applies safely). FCR 1 and FCR 2 incorporate [+REF,+HUM] to 1st and 2nd pronouns; thus (11) (see LEs [15] and [17]) and (12) (see LEs [14] and [24] or [25]) are not LP2 and LP3 acceptable, respectively.

(11) * Pierre le te donne

(12) * Pierre te lui donne

To sum up : observations (30) to (33) in § 1 are accounted for by LP2 to LP5, respectively. Note that GPSG formalism on LP statements enables the formulation of LP 3 with the empirically adequate effect of (31) of § 1.

FCRs 8, 10, 13, and 22 conjointly with ID rules H[1] to H [4] are the main tools handling the morphological variation of pronouns.

[PFORM] is a feature of NPs or of PPs (see FCR 10). It seems that this solution is syntactically motivated (and semantically supported, NPs and PP[PFORMα] being of the same semantic type; see for further discussion A Review of GPSG § 8). If [PFORMα] was exclusively a feature value of PPs, it would be necessary to have H[14'] instead of H[14]:

(13) VP -> H[14'], PP[a, +HUM]

But with (13), [+HUM] ,by HFC, would be transmitted only to the prepositional head and the fact that the corresponding NP must be [+ HUM] would be left unexpressed. PFORM in a NP category functions simply as a morphological marker with

syntactic effects partially expressed by subcategorization frames and partially be LP statements.

ID rules H [1] to H[5] also express the fact that a pronoun to the right of a preposition requires a particular morphological form (e.g. avec toi but not *avec te or *avec tu) and so it is marked[+OBL]. The same feature value is forced onto a topicalized position by FCR 13 (see LE[13]).

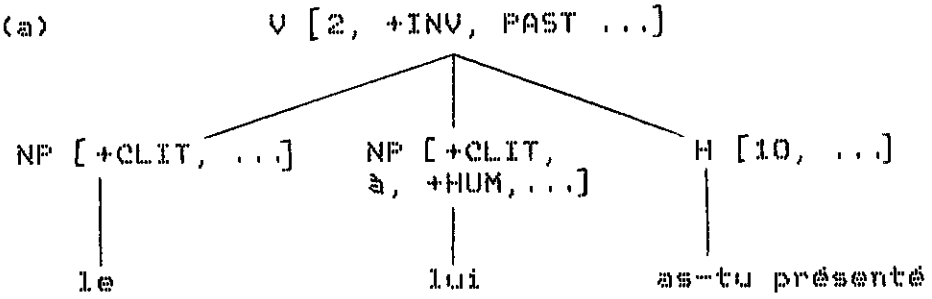
FCR 8 assigns ∼[CAS] to clitic pronouns; their morphology is entirely determined by PFORM, an information carried over in the outputs of the Clitic Metarule (see for example LEs[14] to [16]).

FCR 22 more classically assigns[NOM] to the subject position; (see LE[16]).

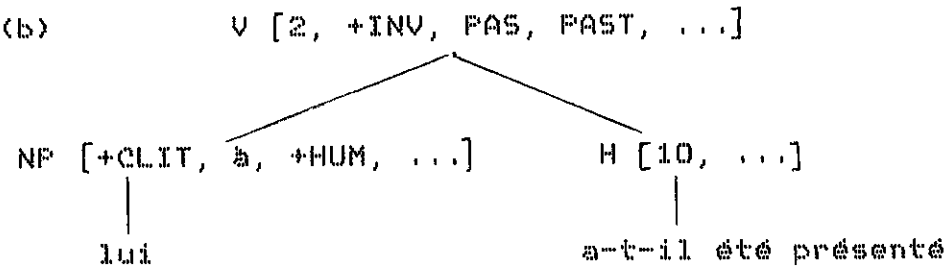
3.3 Clitics and pronoun inversion

CLitics and pronoun inversion can interrelate freely, as required :

(14) (a)



(b)



3.4 Topicalized constructions and wh-questions

The classical GPSG rule for topicalized constructions ($S \rightarrow X^2, H/X^2$), FSD 10 and the (schematic) feature (15)

(15) /[XP, ∼Q]

incorporated to lexical PROFs and verbal entries, are intended to account for the relations between the topicalized

constructions and proforms in the main S. GPSG analysis of unbounded dependency constructions seems thus adequate for handling French relations between topicalized constructions and pronouns (or more generally, proforms), clitic or non clitic.

The value for XP in (15) will be fixed for the different lexical entries. In verbal entries with [+INV], for example, it will always be NP :

(16) (= LE[8])
⟨a-t-il été présenté, [+INV, +PAS, FIN, PAST, AGR NP[PER 3, -PL, MASC], / NP[PER 3, -PL, MASC, ∼Q]] ⟩

In non verbal entries, the slash feature (15) will appear only in PROFs. In pronoun categories, XP will be always NP (see LE[12], [16],[18]). In the other proforms XP = {AP,VP, S} . So for 1(e) we have LE[30] to [32]. Note that for PROFs lexical entries other than PRON there are no parallel entries as LE[17] and LE[18]. The reason is that topicalized non interrogative constituents must be resumed by a PROF but PRON (clitics or non clitics) - contrary to other PROFs - may or may not resume a topicalized constituent.

The feature [∼Q] in (15) precludes an overt proform from resuming an interrogative constituent in topicalized position (see LE[33] and [34]). This construction functions as in GPSG (cf. ST Metarule and LE[35]) but no equivalent of GPSG FCR 22 is incorporated; echo questions in VP are thus specified.

So, as required, it is hoped to attain the following typical results:

(17) (a) G A Que tu connaisses la raison, Pierre le sait
(b) * -A Que tu connaisses la raison, Pierre la sait

(18) (a) G A Intelligent, il l'est
(b) * -A Il l'est

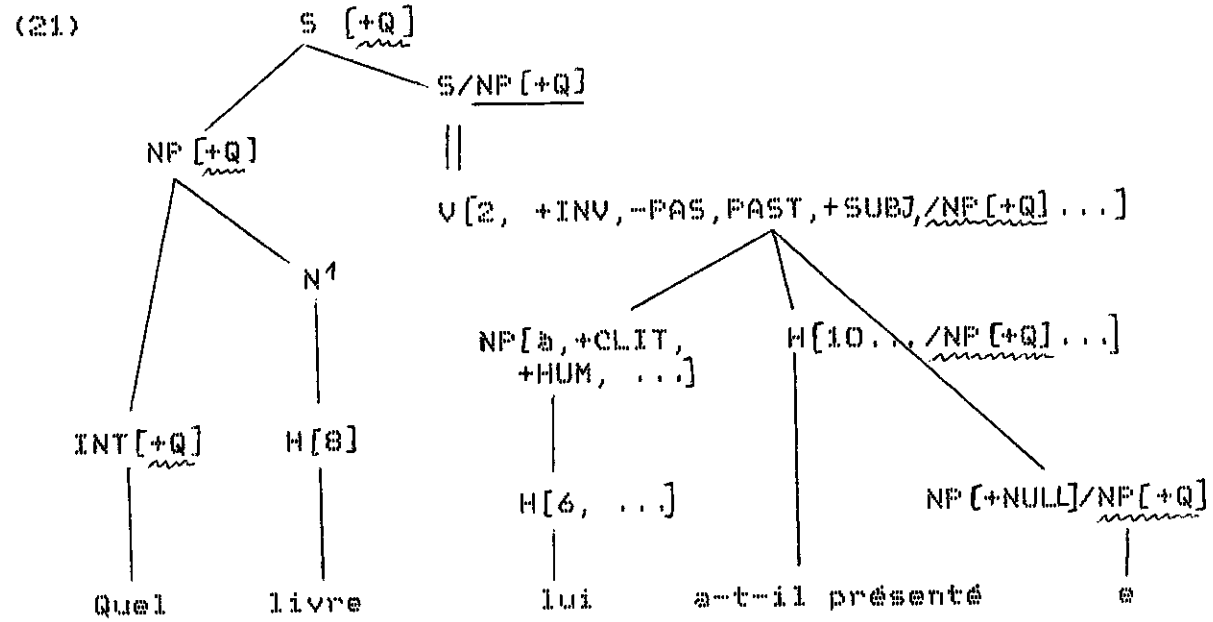
(19) (a) G A Pierre, il présente un cadeau
(b) G A Pierre, Marie lui présente un cadeau
(c) * -A Pierre, Marie présente un cadeau

(20) (a) G A Quel livre Pierre connaît (if FCR 14 is not incorporated)
(b) * -A Quel livre Pierre le connaît

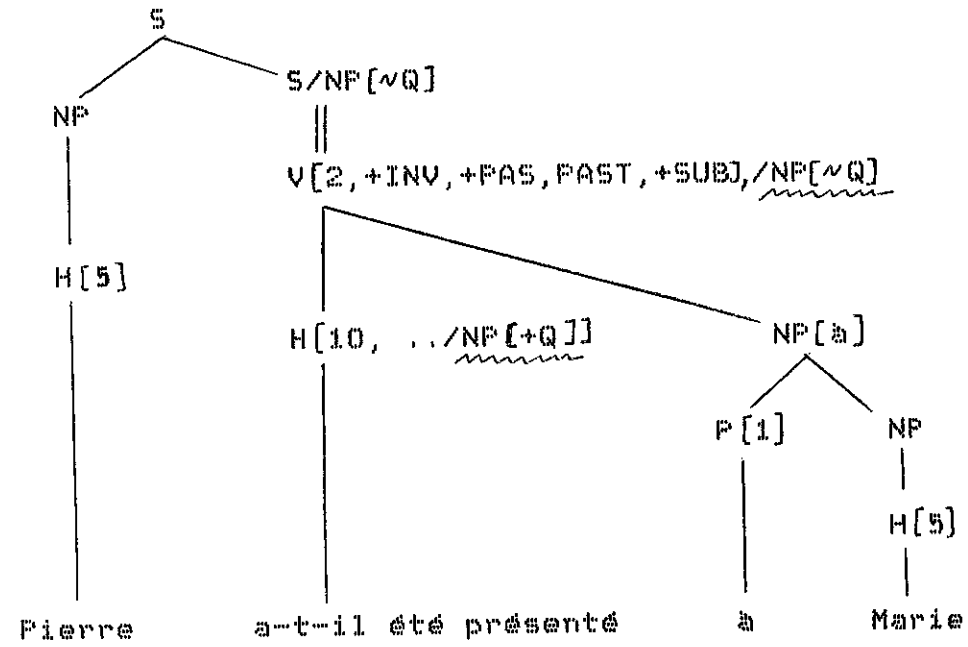
3.5 Topicalized constructions, wh-questions, clitics, inversion and passive.

All the constructions in the title can interact freely in one sentence (but see § 3.6 for further discussion). For example :

(21)



(22)



3.6 Remaining problems and some conclusions

Though the grammar of § 2 is by no means a complete or

definite version (agreement questions are almost entirely neglected), it is intended to cover the observational domains outlined in § 1 with the limits already stated.

It seems clear that GPSG formalism is in principle adequate for specifying, along the lines of § 2, a French grammar covering such constructions as clitics, topicalization, pronoun inversion and passive. Note that no important change to the core of GPSG formalism has been proposed; in particular, the universal feature instantiation principles remain unchanged.

The proposed grammar can be improved by adding more details. In particular, this is true for lexical entries, where much is to be done if it is desired to approach true effectiveness. Furthermore, more constructions (e.g. dans NP resumed by y) can be covered along the same lines. It seems also that the already pointed out limitations (for example, about se) can be overcome in the same way. Finally, it is hoped that simplifications can be introduced. It was already noted that an analysis altogether eliminating the Pronoun Inversion Metarule is quite workable. The same seems to be true for the Clitic Metarule : its only effect is to incorporate a feature value. But this can be left as a free instantiation one. The correct placement of the resultant clitic categories will be reserved for the LP statements alone. In regard to clitics, it will thus be possible to only incorporate in the grammar metarules within the constraint that a metarule must not have as input an ID rule which is the output of the application of this same metarule.

But there are other French constructions involving clitics and topicalized constructions which deserve much more careful attention.

There are well attested constructions with clitic en that lie outside the limits of the proposed grammar. Among others :

(23) (a) Pierre court après Marie
(b) Pierre lui court après

(24) (a) Pierre a pris quatre livres
(b) Pierre en a pris quatre

(25) (a) Il est conscient de cela
(b) Il en est conscient

The trouble with (23) to (25) is that they involve the cliticization of a constituent within a constituent.

On the other hand, it can be observed that all samples already presented involve only one constituent in topicalized position. But a plurality of them is quite possible :

(26) Quel livre Pierre veut-il lire?
Which book Peter does he want to read?

(27) Pierre, ce livre il l'aime
Peter, this book he likes it

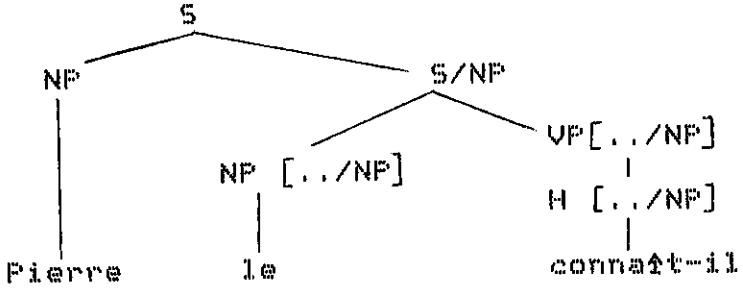
(28) Que tu connaisses le problème, Pierre il le sait
That you know the problem, Peter, he knows it

(29) Marie, intelligente elle l'est
Mary, intelligent she is

It seems that the GPSG formalism must be adapted to handle these constructions. It is possible to think of at least two ways for doing so : (a) a category value feature may be allowed to have more than one category as value (see GPSG p. 81), or (b) there may be more than one slash feature (say, "/", "//", etc.) each one with only one C as value.

Moreover, the grammar of § 2 presents inadequacies that are the analog of inadequacies pointed out for GPSG English grammar (see above "A review of GPSG" § 7). The following result is indeed obtained :

(30) (a) Pierre il le connaît
(b)



This is disturbing in as much as it may be wanted to interrelate the information that FFP and SLASH can transmit from one point of the tree to another with anaphora and ellipsis issues. (30 a) does not mean (31); but compare instead (30 a) with (32 a).

(31) Pierre se connaît-il?
Peter does he know himself?

(32) (a) Pierre Marie le connaît
Peter Mary knows Peter

(b) Pierre il la connaît
Peter he knows her

These issues will be reconsidered in the more general conclusion of this report.

SOME REMARKS ON INFINITIVAL VP COMPLEMENTS IN FRENCH

The way they do account for Control phenomena is a main topic in evaluating grammars. Thus we examine here how it is possible to adapt GPSG treatment of Control to French.

1. Lexical ID rules

We list below most of the lexical ID rules which are needed for infinitival-complement constructions in French. They account for almost all the 2800 verbs described in Gross (1975), therefore their great number.

There is no French correspondent to the English 'VP -> to VP' (i.e. VP[INF, +AUX] -> H[12], VP[BSE]) rule, instead we use the following one :

(1) $VP\ [PREP\ \alpha] \rightarrow \{[SUBCAT\alpha]\}, H$

where $\alpha \in \{$ à, de, pour $\}$

And according to facts about clitics, we rewrite the GPSG PP [PFORM β] as NP[β], where NP[β] stands for NP [PFORM β] (see Bès for discussion). So we assume in addition :

(2) $NP\ [PFORM\ \beta] \rightarrow P[n]\ ,\ H$

where $\beta \in \{$à, de, pour, avec, sur ...$\}$

The rules are :

VP -> H[1], VP[INF]
  oser, croire, aimer, ...
  ose rêver.

VP -> H[2], VP[INF, PREP à]
  chercher, commencer, demander, ...
  cherche à partir.

VP -> H[3], VP[INF, PREP de]
  oublier, tenter, craindre, ...
  oublie de venir.

VP -> H[4], (PP[+LOC]), VP[INF]
  aller, courir, descendre, ...
  descend (dans la rue) chercher son livre
    (chez Marie)

VP -> H[5], PP[+LOC], VP[INF]
  s'engouffrer, naviguer, couler, ...
  s'engouffre dans le couloir retrouver Marie

```
VP  ->  H [6], (NP[+LOC]), VP[INF]
        monter, descendre, grimper, ...
        monte (l'escalier) rejoindre Marie

VP  ->  H [7],  NP[+LOC], VP[INF]
        dévaler, franchir, traverser, ...
        dévale les marches rejoindre Marie

VP  ->  H [8], (NP), VP[INF]
        envoyer, entendre, laisser, ...
        envoie (Marie) chercher Pierre

VP  ->  H [9], NP, VP[INF]
        conduire, amener, imaginer, ...
        conduit Marie chercher Pierre

VP  ->  H[10], NP, VP[INF, PREP à]
        décider, obliger, motiver, ...
        décide Marie à travailler

VP  ->  H[11], NP[+ABS], VP[INF, PREP à]
        dépenser, passer, perdre, ...
        dépense une grosse somme à acheter cela

VP  ->  H[12], NP, VP[INF, PREP de]
        empêcher, persuader, aimer, ...
        empêche Marie de partir

VP  ->  H[13], NP, VP[INF, PREP de]
        menacer, persuader, convaincre, ...
        menace Marie de partir seul

VP  ->  H[14], NP, VP[INF, PREP pour]
        embaucher, convoquer, engager, ...
        embauche Pierre pour finir ce travail

VP  ->  H[15], NP[à], VP[INF, PREP à]
        apprendre, montrer, servir, ...
        apprend à Marie à lire

VP  ->  H[16],  NP[à],  VP[INF,  PREP de]
        souhaiter, éviter, refuser,
        souhaite à Marie de réussir

VP  ->  H[17], NP[pour], VP[INF, PREP à]
        aboutir, consister, servir, ...
        aboutit pour Marie à devoir partir

VP  ->  H[18], NP[pour], VP[INF, PREP de]
        exclure, exiger, prévoir, ...
        exclut pour Marie de partir seule

VP  ->  H[19], NP[de], VP[INF, PREP de]
        exiger, obtenir, ...
        exige de Marie de partir seule
```

```
VP  ->  H[20], (PP), VP[INF, PREP à]
        s'accorder, collaborer, appprendre, ...
        s'accorde (avec Marie) à finir ce travail

VP  ->  H[21], (PP), VP[INF, PREP de]
        manigancer, parier, tenir, ...
        manigance (avec Marie) de partir seul

VP  ->  H[22], (PP), VP[INF, PREP pour]
        s'arranger, s'entendre, ...
        s'arrange (avec Marie) pour partir seul

VP  ->  H[23], (NP[à]), VP[INF, PREP de]
        ordonner, dire, proposer, ...
        ordonne (à Marie) de partir

VP  ->  H[24], (NP[à]), VP[INF, PREP de]
        promettre, proposer, ...
        promettre (à Marie) de partir

VP  ->  H[25], (NP[à]), VP[INF]
        dire, sembler, ...
        dit (à Marie) être parti seul

VP[+il]  ->  H[26], VP[INF]
        sembler, paraître, ...
        semble pleuvoir

VP[+il]  ->  H[27], (NP[à]), VP[INF]
        falloir, ...
        faut (à Marie) travailler beaucoup

VP[+il]  ->  H[28], NP[à], VP[INF]
        sembler, ...
        semble (à Marie) avoir travaillé beaucoup

VP[+il]  ->  H[29], NP[à], VP[INF, PREP de]
        souvenir, ...
        souvient à Marie d'être allée là-bas

VP[+il]  ->  H[30], (NP[pour]), VP[INF]
        faire bon, ...
        fait bon (pour Marie) se reposer

VP[+il]  ->  H[31], (NP[pour]), VP[INF, PREP à]
        y avoir avantage, ...
        y a avantage (pour Marie) à travailler beaucoup

VP[+il]  ->  H[32], (NP[pour] ), VP[INF, PREP de]
        s'agir, être besoin, être l'heure, ...
        s'agit (pour Marie) de se reposer un peu
```

2. LP statements

Word order in French seems to be freer than in English. To account for

(3) Jean aime Marie
* Aime Marie Jean
* Jean Marie aime

we assume the following LP statements :

[SUBCAT α, -CLIT] < [$\sim$SUBCAT, -CLIT]

NP < VP[$\sim$INF]

Since we have

(4) Jean va chez Marie chercher son livre

(5) Jean ordonne à Marie de partir

(6) Jean entend Marie chanter

as well as

(7) Jean va chercher son livre chez Marie

(8) Jean ordonne de partir à Marie

(9) Jean entend chanter Marie

we do not constrain the precedence relations between NPs, PPs and VPs [INF]. However, they have to be constrained in some particular cases :

(10) Jean monte l'escalier rejoindre Marie
* Jean monte rejoindre Marie l'escalier

(11) Jean envoie Marie chercher Pierre
* Jean envoie chercher Pierre Marie

(12) Jean décide Marie à travailler son anglais
* Jean décide à travailler son anglais Marie

(13) Jean hasarde sa vie à entreprendre cela
* Jean hasarde à entreprendre cela sa vie

If we can, for example (10) (i.e. for V[4], V[5], V[6], V[7],) state something like

XP [+LOC] < VP[INF]

we cannot state

NP[~PFORM] < VP[INF]

to handle (11)-(13), because we want to allow (9) and also the following sentences :

(14) Jean décide à travailler son anglais la soeur de Marie
Jean hasarde à entreprendre cela sa vie toute entière
Jean a persuadé d'être le seul compétent le chef du service du personnel
etc.

Acceptability is here clearly tied to the length of NP and VP INF constituents : the problematic question that then arises is how to make LP statements "context-sensitive" in such a way.

3. Obligatory Control

3.1. If, as in GPSG, we take the basic semantic type of a verb to be the type the verb has when it is combined with a that-clause complement, there are two cases where the semantic S-analysis has no syntactic correspondence :

a) The verb admits no clausal-complement at all :

(15) Jean ose partir
* Jean ose qu'il part/qu'il parte

Jean commence à travailler
* Jean commence (à ce) qu'il travaille

b) The verb admits a clausal-complement, but the relation between the infinitival and the clausal-complement is not one of paraphrase :

(16) $Jean_1$ veut $_1$partir
$Jean_1$ veut qu'il_2 parte/* qu'il_1 parte

$Jean_1$ cherche à $_1$partir
$Jean_1$ cherche à ce qu'il_2 parte/* à ce qu'il_1 parte

But the assumption that all control verbs are analyzed in terms of basic expressions which take propositional arguments allows to capture the notion of "understood subject" : in case (a) we obtain the interpretation we would have if there was a clausal-complement with (among others) a coreferential reading, and in case (b) we obtain the one we would have if the clausal-complement had (among others) a coreferential reading. Actual paraphrase relations, such as

between (17) and (18).

(17) Jean$_1$ ordonne à Pierre$_2$ qu'il$_2$ parte

(18) Jean$_1$ ordonne à Pierre$_2$ de $_2$partir

are established by translating both ordonner que and ordonner de by means of the same IL constant (say, ordonner').

We leave open the question whether the two semantic combinators f_E and f_R are really necessary or not (for discussion, see Klein & Sag (1985) and Dowty (1985)), and use them to build the relevant types of infinitival-complement taking verbs in French. This is illustrated in the following table :

SUBCAT number	VERB	TRANSLATION	TYPE OF α'	RESULTING TYPE
1	oser	f_E (oser')	⟨S,⟨NP,S⟩⟩	⟨VP,⟨NP,S⟩⟩
2	chercher	f_E (chercher')	⟨S,⟨NP,S⟩⟩	⟨VP,⟨NP,S⟩⟩
3	oublier	f_E (oublier')	⟨S,⟨NP,S⟩⟩	⟨VP,⟨NP,S⟩⟩
* 4	aller	$\lambda \mathcal{P} f_E$ (aller' ($\mathcal{P}$))	⟨PP,⟨S,⟨NP,S⟩⟩⟩	⟨PP,⟨VP,⟨NP,S⟩⟩⟩
* 5	s'engouffrer	$\lambda \mathcal{P} f_E$ (s'engouffrer' ($\mathcal{P}$))	⟨PP,⟨S,⟨NP,S⟩⟩⟩	⟨PP,⟨VP,⟨NP,S⟩⟩⟩
6	monter	$\lambda \mathcal{P} f_E$ (monter' ($\mathcal{P}$))	⟨NP,⟨S,⟨NP,S⟩⟩⟩	⟨NP,⟨VP,⟨NP,S⟩⟩⟩
7	dévaler	"	"	"
8	envoyer	f_R (envoyer')	⟨S,⟨NP,S⟩⟩	⟨VP,⟨NP,⟨NP,S⟩⟩⟩
9	conduire	"	"	"
10	décider	f_E (décider')	⟨S,⟨NP,⟨NP,S⟩⟩⟩	⟨VP,⟨NP,⟨NP,S⟩⟩⟩
11	dépenser	$\lambda \mathcal{P} f_E$ (dépenser' ($\mathcal{P}$))	⟨NP,⟨S,⟨NP,S⟩⟩⟩	⟨NP,⟨VP,⟨NP,S⟩⟩⟩
12	empêcher	f_R (empêcher')	⟨S,⟨NP,S⟩⟩	⟨VP,⟨NP,⟨NP,S⟩⟩⟩
	persuader	f_E (persuader')	⟨S,⟨NP,⟨NP,S⟩⟩⟩	⟨VP,⟨NP,⟨NP,S⟩⟩⟩
13	menacer	$\lambda \mathcal{P} f_E$ (menacer' ($\mathcal{P}$))	⟨NP,⟨S,⟨NP,S⟩⟩⟩	⟨NP,⟨VP,⟨NP,S⟩⟩⟩
14	embaucher	f_E (embaucher')	⟨S,⟨NP,⟨NP,S⟩⟩⟩	⟨VP,⟨NP,⟨NP,S⟩⟩⟩
15	apprendre	"	"	"
16	souhaiter	"	"	"
17	aboutir	"	"	"
18	exclure	"	"	"
19	exiger	"	"	"
*20	s'accorder	$\lambda \mathcal{P} f_E$ (s'accorder' ($\mathcal{P}$))	⟨PP,⟨S,⟨NP,S⟩⟩⟩	⟨PP,⟨VP,⟨NP,S⟩⟩⟩
21	manigancer	"	"	"
22	s'arranger	"	"	"
23	ordonner	f_E (ordonner')	⟨S,⟨NP,⟨NP,S⟩⟩⟩	⟨VP,⟨NP,⟨NP,S⟩⟩⟩
24	promettre	$\lambda \mathcal{P} f_E$ (promettre' ($\mathcal{P}$))	⟨NP,⟨S,⟨NP,S⟩⟩⟩	⟨NP,⟨VP,⟨NP,S⟩⟩⟩
25	dire	"	"	"

* The trouble here is that $\mathcal{P}$ stands in principle for NP-types, and PPs are not of type NP.

Given these type-assignments, we can assume the GPSG definition of control (p. 203) :

> Let α , β and γ be constituents whose types are $\langle VP, \langle NP, a\rangle\rangle$, VP and NP respectively, and whose translations are α', β', and γ'.
> If $0 = \alpha'(\beta')(\gamma')$, then γ controls β .

to yield for example correctly :

(19) ordonne' (partir') (Marie*) (Jean*)
Jean$_1$ ordonne à Marie$_2$ de $_2$partir

(20) promet' (Marie*) (partir') (Jean*)
Jean$_1$ promet à Marie$_2$ de $_1$partir

Correct results are also obtained for all the other verbs.

3.2 We sometimes found a syntactically-marked motivation for assigning $\langle NP, \langle VP, \langle NP, S\rangle\rangle\rangle$ instead of $\langle VP, \langle NP, \langle NP, S\rangle\rangle\rangle$, that is, a non ad hoc relation between syntax and semantic types. We leave open the question whether the GPSG formalism can express this kind of generalization. We have for example :

a) A PP is not a possible controller

b) An inherited NP[+LOC] (i.e. l'escalier, la rue, la cour ...) or an inherited NP[+ABS] (for 'abstract', i.e. une grosse somme, du temps, sa vie, son énergie, ses forces ...) are not possible controllers.

c) For an NP to be a controller, the infinitival complement must be VP[INF, PREPα] when the head verb admits both an NP and an S argument.

Suppose the grammar to embody among the ID rules the following:

(21) VP -> H[n], (NP[à]), S[FIN]
dire, écrire, rappeler, ...

What (c) says is that Marie can be the controller of VP[INF] in

(22) Jean$_1$ dit à Marie$_2$ de $_2$partir seule

but not in

(23) Jean$_1$ dit à Marie$_2$ $_1$être parti seul.

And (c) is not true for

Jean$_1$ entend Marie$_2$ $_2$chanter

Jean$_1$ imagine Marie$_2$ $_2$chanter

because the corresponding rule is

(24) VP -> H[m], S[FIN]

(25) Jean entend que Marie chante
* Jean entend Marie qu'elle chante
Jean imagine que Marie chante
* Jean imagine Marie qu'elle chante

In other words, (c) counts for Equi, not for Raising predicates.

3.3 On the other hand, there seems to be a systematic correlation between clausal and infinitival complements. According to whether the former takes indicative and/or subjunctive, we may infer which sentence the latter does in fact paraphase. Consider the examples :

(26) Jean$_1$ ordonne à Pierre$_2$ qu'il$_2$ parte

Jean$_1$ ordonne à Pierre$_2$ de $_2$partir

* Jean$_1$ ordonne à Pierre$_2$ qu'il$_1$ parte

* Jean$_1$ ordonne à Pierre$_2$ de $_1$partir

(27) Jean$_1$ propose à Pierre$_2$ qu'il$_2$ parte le premier

Jean$_1$ propose à Pierre$_2$ de $_2$partir le premier

Jean$_1$ propose à Pierre$_2$ qu'il$_1$ parte le premier

Jean$_1$ propose à Pierre$_2$ de $_1$partir le premier

In some cases (cf. (26) vs. (27))the subjunctive forbids coreference between matrix and embedded subjects. Such a prohibition generally remains in corresponding infinitives, as inspection of the following Equi-predicates shows :

	S indicative	S subjunctive	VP -> H[n], (XP), S[FIN]	Coreference forbidden between the matrix subject and the subject of S subj.	Subject-Controlled Equi Verb	Object-Controlled Equi Verb	Infinitive paraphases ...
demander [2] souhaiter [1]	-	+	-	-	+	-	S subj.
vouloir [1] aimer [1]	-	+	-	+	+	-	S subj.
V [14] -> V [19] ordonner [23]	-	+	+	+	-	+	S subj.
proposer	-	+	+	-	+ (proposer [24])	+ (proposer [23])	S subj.
croire [1]	+	-	-	-	+	-	S ind.
promettre [24]	+	-	+	-	+	-	S ind.
décider [3]	+	+	-	+	+	-	S ind.
admettre	+	+	-	-	+	-	S ind. admettre [1] S subj. admettre [3]
dire	+	+	+	+	+	-	S ind. dire [25]
					-	+	S subj. dire [23]

Thus we can state that

a) relations of coreference forbidden by means of S subjunctive remain forbidden in the corresponding infinitival construction, i.e. there is no Subject-Controlled-Equi predicate which paraphrases S subjunctive-sentences. (The only exception is the case where we have the one and only possibility of an S-subjunctive sentence; here the Subject-Controlled-Equi predicate is obligatory, but it paraphrases no actual S-subjunctive sentence :

(28) * Il_1 veut qu'il_1 vienne

Il_1 veut qu'il_2 vienne

* Il_1 veut qu'il_1 viendra

* Il_1 veut qu'il_2 viendra).

b) when there exists no prohibition on coreference, we have Object-Controlled-Equi predicates (if the NP-object is a possible controller) and Subject-Controlled-Equi predicates.

(Work in progress will systematize such observations in order to make the lexicon consistent with appropriate semantic entailments).

3.4 To handle optional arguments, it is possible to adopt a principle which ensures existential quantification into missing argument positions; see for example Gazdar (1982) Optional Argument Convention (p. 151) :

> If β ' = ...(α") ..., where β immediately dominates an optional constituent α and α "is of type $\langle s,\langle\langle e,t\rangle,t\rangle\rangle$, then when α is omitted, β ' = ... $(\hat{P}[\exists x \ P(x)])$.

Thus the interpretation of

(29) Jean ordonne de partir

is something like

(30) $(\exists x)$ (Jean ordonne à x de partir)

where x controls the infinitival VP.

But if there is no controller at all, including the case where the matrix subject is the dummy NP il (which translates by means of vacuous-lambda abstraction), arbitrary reference may introduce universal quantification :

(31) Partir, c'est mourir un peu
$(\forall x)$ (x partir = x mourir un peu)

Il faut manger pour vivre
$(\forall x)$ (il faut x manger pour x vivre)

This could be done by formulating a 'No Controller' Convention.

3.5 Semantic interpretation should sometimes allow for more than one controller at once :

(32) Jean$_1$ manigance avec Pierre$_2$ de $_{1+2}$ partir les premiers

Jean$_1$ propose$_2$ à Pierre de $_{1+2}$ partir les premiers

That is a problematic topic because it seems very difficult to express this possibility within semantic types (in a way which does not contradict the above definition of Control).

4. Infinitive VPs in Prepositional Phrases

If we suppose that the grammar includes rules (33) and (34) :

(33) S -> NP, VP, $(V^2$ [+FINAL])

(34) VP[+FINAL] -> PREP γ , VP[INF]

where $\gamma \in \{$ pour, afin de, ... $\}$

it is possible to extend the control process to the following sentences (not treated in GPSG) :

(35) Jean$_1$ travaille pour $_1$réussir

(36) Jean$_1$ veut $_1$travailler pour $_1$réussir

Taking VP [+FINAL] to be of type $\langle$VP,VP$\rangle$, functional realization will build a complex VP (pour réussir (travailler)) which undergoes clause (i) of the general definition of control (GPSG chapter 5, p. 88, def. 4) :

If $\emptyset$ is a projection of r, where r = C_0 -> C_1, ..., C_n, then a category $\emptyset(Ci)$ controls $\emptyset$ (Cj) in $\emptyset$, $1 \leq i,j \leq n$, iff (i) TYP $(\chi(\emptyset(Cj))) = \langle$TYP$(\chi(\emptyset(Ci)))$, TYP$(\chi(\emptyset(Co)))$, (...).

Thus Jean controls the whole VP travaille pour réussir in (35) and veut travailler pour réussir in (36).

Given the LP statements assumed above, the grammar admits :

(37) Jean$_1$ travaille ses maths pour $_1$réussir

Jean$_1$ veut $_1$travailler ses maths pour $_1$réussir

(38) Jean$_1$, pour $_1$réussir, travaille ses maths

Jean$_1$, pour $_1$réussir, veut $_1$travailler ses maths

(39) Pour $_1$réussir, Jean$_1$ travaille ses maths

Pour $_1$réussir, Jean veut $_1$travailler ses maths

but not :

(40) ?* Jean$_1$ travaille, pour $_1$réussir, ses maths

?* Jean$_1$ veut $_1$travailler, pour $_1$réussir, ses maths

(41) Jean$_1$ veut, pour $_1$réussir, $_1$travailler ses maths

The trouble here is that (41) is grammatical and that (40) becomes acceptable with an NP object of greater length :

(42) Jean$_1$ travaille, pour $_1$réussir, ses maths et son anglais

Jean$_1$ veut $_1$travailler, pour $_1$réussir, ses maths et son anglais

(see paragraph 2).

Moreover, we posit the following Metarule

(43) VP $\rightarrow$ W, VP[INF]

$\Downarrow$

VP $\rightarrow$ W, VP[INF], V^2 [+FINAL]

The output of (43) for, say V[23], will be rule (44) :

(44) VP $\rightarrow$ H[23], (NP[à]), VP[INF, PREP de], V^2 [+FINAL]

and the semantic type associated with its head (taking V^2 to be VP, not S, and the semantic type of VP[+FINAL] to be ⟨VP,VP⟩ again) :

(45) ⟨⟨VP,VP⟩, ⟨VP,⟨NP,VP⟩⟩⟩

The complex VP formed by combining VP[INF] and VP [+FINAL] will

therefore undergo the same control as a "bare" VP[INF] . We account now for the examples :

(46) Jean$_1$ ordonne à Marie$_2$ de $_2$travailler beaucoup pour $_2$réussir ses examens

Jean$_1$ ordonne à Marie$_2$, pour $_2$réussir ses examens, de $_2$travailler beaucoup

Jean$_1$ ordonne de $_2$travailler beaucoup à Marie$_2$, pour $_2$réussir ses examens

Note that, according to LP statement

NP < VP[$\sim$INF]

we correctly rule out sentences like

(47) * Jean$_1$ ordonne, pour $_2$réussir ses examens, à (la soeur de) Marie$_2$ de $_2$travailler beaucoup

* Jean$_1$ ordonne, pour $_2$réussir ses examens, de $_2$travailler beaucoup à (la soeur de) Marie$_2$

?* Jean$_1$ ordonne de $_2$travailler beaucoup, pour $_2$réussir ses examens, à (la soeur de) Marie$_2$

And since the rule involving VP[+FINAL] is here (44), not (33), we also block :

(48) * Pour $_2$réussir ses examens, Jean$_1$ ordonne à Marie$_2$ de $_2$travailler

* Jean$_1$, pour $_2$réussir ses examens, ordonne à Marie$_2$ de $_2$travailler

With both rules (33) and (44) we obtain :

(49) Pour $_1$l'aider, Jean$_1$ ordonne à Marie$_2$ de $_2$travailler beaucoup pour $_2$réussir ses examens

Jean$_1$, pour $_1$l'aider, ordonne à Marie$_2$ de $_2$travailler beaucoup pour $_2$réussir ses examens

Jean$_1$ ordonne à Marie$_2$ de $_2$travailler beaucoup pour $_2$réussir ses examens, pour $_1$l'aider

An obvious consequence of such a treatment is that the sentence

(50) Jean$_1$ promet à Marie$_2$ de $_1$travailler beaucoup pour $_1$réussir dans son projet

will be taken as being ambiguous (whether analyzable by means of (33) or (44)). This does not seem unreasonable; we

likewise agree with the assumption that

(51) Jean$_1$ veut $_1$travailler pour $_1$réussir

is ambiguous with respect to the two (informally written) following meanings :

(52) (veut travailler) (pour réussir) [= (33)]

(53) (veut) (travailler pour réussir) [= (44)]

(Note again that (44) allows sentence (41) which we could not admit by virtue of (33)). There is perhaps more evidence in case (50) :

(54) (promet de travailler) (pour réussir) [= (33)]

(55) (promet) (de travailler pour réussir) [= (44)]

where we understand "success" to be closely tied either to "promise" or to "work", respectively.

5. Conclusion

According to this first investigation, it seems that the infinitival constructions whose English counterparts are accounted for in GPSG may find an adequate treatment in French.

However we have to take care of Linear Precedence, a matter which remains puzzling. GPSG control is also seemingly relevant for enlargements such as in paragraph 4, but increasing the data in order to provide the grammar with, say, indirect infinitival interrogatives, infinitival relatives, clitics of infinitival VP complements, adverbial infinitives and so on, will at least lead us to add new assumptions to the previous definitions. Consider for example

(56) $_1$Travailler ennuie Marie$_1$

i.e. a sentence where an infinitive occupies the subject position. At present there is no way to ensure that a NP object has to control something in subject position : but further research into such questions will precisely put to the test the formal possibilities of GPSG theoretical framework.

REFERENCES

DOWTY (1985), On Recent Analyses of the Semantics of Control, Linguistics and Philosophy 8, p. 291-331.

GAZDAR (1982), Phrase Structure Grammar, in P. Jacobson and G.K. Pullum (eds.), The Nature of Syntactic Representation, D. Reidel, p. 131-186.

GROSS (1975), Méthodes en Syntaxe, Hermann, Paris.

KLEIN & SAG (1985), Type-Driven Translation, Linguistics and Philosophy 8, p. 163-201.

CONCLUSION AND SUGGESTIONS

1. The following main points resume this report on the feasibility of a GPSG French grammar.

1.1 GPSG is a very explicite clear and neat model. Though a sophisticated one, its descriptive constraints are formulated in a way that revision of particular analysis with their consequences is feasible (perhaps even calculable). Nevertheless, some more explication on several points will be welcome. ECPO (see p. 3), FSD and admissibility (see p. 4), lexical entries (see p. 9) and metarules (see p. 15).

1.2 Within the limits of the revised constructions, no unsurmontable descriptive problems have been detected. The descriptive solutions and suggestions proposed throughout this report must be considered as a support of the claim that problems are workable within a GPSG formalism.

Two (minor) points remain :

(a) The introduction of morphological contextual features (e.g. l'élire and not *le élire).

(b) The formulation of LP statements in function of the length of the constituents (see p. 44).

1.3 The central features of the GPSG syntactic formalism may remain unchanged. The universal instantion principles work without major problems in the French grammar.

The issue with CAP (see p. 10) may be overcome by reworking the analysis on CASE. The one concerning SLASH and FFP (see p. 40) requires adjustments but no radical changes.

1.4 French constituent order is rather puzzling : sometimes very rigid, elsewhere quite free. The dissociation of order relations from dominance relations in GPSG formalism, seems to be a powerful tool enabling to handle many French problems. Note that in addition to the preceding discussions (see p. 35 and p. 44), the suggested LP statements account correctly for

(1) (a) Pierre, à la poste, il y va régulièrement
(b) A la poste, Pierre il y va régulièrement

(2) (a) Pierre tu le connais
(b) Tu le connais, Pierre

1.5 But it is also true that some rather clear

generalizations are left unexpressed.

1.5.1 Infinitival VP can be cliticized and the clitic form is determined not by the ID lexical rule which has the VP[INF] category as a sister of the verbal head, but by the sisters of a verbal head of another ID lexical rule of the "same" verb.

Example : oublier et désespérer are both H[3](see p. 41), but we have

(3) (a) (D') envoyer la lettre, je l'ai oublié
(b) Je l'ai oublié, d'envoyer la lettre

(4) (a) *D'envoyer la lettre, j'en ai oublié
(b) *J'en ai oublié, d'envoyer la lettre

(5) (a) De partir ce soir, j'en désespère
(b) J'en désespère, de partir ce soir

(6) (a) *De partir ce soir, je le désespère
(b) *Je le désespère, de partir ce soir

(7) VP -> H[n], NP (oublier)

(8) VP -> H[n], NP[de] (désespérer)

Clitic forms in (3) and (5) are thus determined by (7) and (8) respectively, and not by H[3].

1.5.2 It is possible to make analogous remarks on the semantic types which must be associated with lexical entries (see p. 48). The specifications of a plurality of controllers remain open (see p. 52).

1.6 An extended observational domain must be treated (see below § 2) before concluding. It seems clear that the GPSG model has not yet been pushed to its limits. Among others, the following questions ought to be treated : (a) agreement; (b) anaphora and ellipsis; coreference and reflexives; (c) coordination (it may be that the proposed analysis of complex verb forms works with some difficulty here (cf. avoir acheté et vendu des livres); (d) infinitive constructions not embedded in VPs; (e) tense and other verbal categories; (f) an extended NP.

2. Suggestions

Many points have to be worked on before attaining an effective solution to the question - answer problem. We suggest that it would be preferable to concentrate firstly on

the simplest form of a dialogue; i.e. one question and one answer. The goal can be to try to approach exhaustivity of interrogative constructions within these limits.

We also suggest that comparisons of linguistic models can be enriched by the confrontation of the empirical adequacy of specified particular grammars, if constraints on the enlargement of the observational domains are accepted. It may be that the merits of one model compared to another ought to be reconsidered if detailed grammars covering much more than some classical examples are compared.

But to do so, it appears that it is necessary to dispose of some vocabulary which must enable the specification of observational domains. This vocabulary must be as "neutral" as possible i.e. not conditioned by the different models to be tested; conceptually, it will work in a way close to ostentive definitions. With a tool of this sort, the comparison of the (sub) languages that each grammar can adequately specify will be easier than at present. Empirical adequacy seems an unescapable parameter (though not a sufficient one) in the comparison of competing linguistic models.

FURTHER COMMENTS

Thanks to Klein's comments (Notes of January 22 1986) and to a long discussion at Marcoussis with Annick Corluy and Thierry Guillotin, it was possible to introduce some modifications (rectifications and/or enlightenments) to the primitive text (see p. 9, 10, 15, 21, 48, 49). Nevertheless, all the received observations were not incorporated : questions related to borderline phenomena (grammatical judgements, limits of syntax, etc.) were not modified. In any case, we agree that some of our observations are no more founded if the grammar is to be supplemented with a theory of speech acts (see above §3.2 p. 5 and (1) §7.1 p. 14).

It is now possible to sum up the crucial questions which must be further studied:

(a) General constraints in the specification of lexical entries (see §5 p. 9) The relation between a lexical head node and its leaf is not transparent. GPSP proposition in p. 34 does not seem to work well with the suggestion in note 20 p. 107 if further conventions are not added. In the proposed GPSG French grammar, two different notations were introduced for handling essentially the same problem (LE 1 to 3 with GENDER, and LE 24 and LE 25 for lui). We feel that Klein's suggestions (§5 of his comments) are profitable general ideas on which to work.

(b) CAP must be reformulated in general grounds and agreement French questions carefully studied.

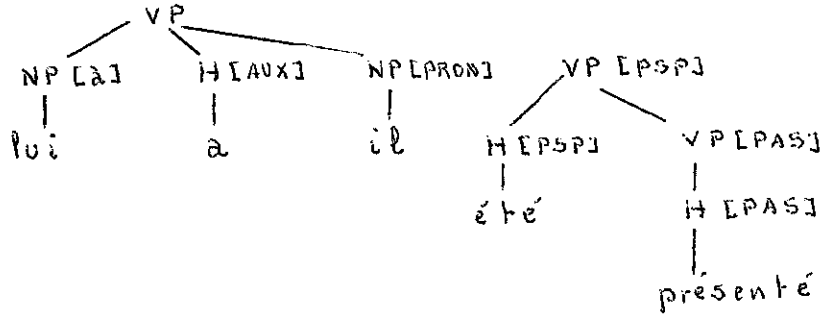
(c) Some problems of unbounded dependencies remain : see above §6.2 p.10 and some if not all questions raised in §7.

(d) We think that there is a difference in principle between (1) and (2): English admits (3) and (4) but there are no corresponding French sentences (5) and (6) :

(1) Has he been introduced?
(2) A-t-il été présenté?
(3) Has the boy been introduced?
(4) Has the boy who was there been introduced?
(5) * A le garçon été présenté?
(6) * A le garçon qui était là était présenté?

So the proposed treatment of VPs is possible for French but impossible for English. In this first and experimental GPSG French grammar, the strongest motivation for the proposed treatment was the handling of clitic 'movement'. Assuming (7) we run into trouble in a more orthodox GPSG approach to auxiliaries and passive to relate it to the local tree of type (9) which must be assoted with (8):

(7) VP → H[n] NP NP[à]
(8) Lui a-t-il été présenté?
(9)



The trouble is the same with an analysis where only forms such as a-t-il are listed.

But saying that the proposed analysis is a possible one does not mean that it is a good one. As indicated, the proposed GPSG French grammar does not cover all problems related to clitics and much is to be done if some degree of effectiveness is desired. It seems rather possible that with a revised and more general treatment of clitics, a new and more elegant analysis of VPs can be be worked out.

(e) We think nevertheless that the specification of UCG modifies work priorities. It seems rather urgent now to study the new model and to apply it in some experimental manner to French before pursuing the task of implementing a GPSG French grammar. Though UCG in several important points makes different choices than GPSG, there seem to be many fundamental common abstract features between GPSG and UCG. So it is thought that some pending problems in GPSG (say agreement, specification of lexical entries, and, as pointed out by Klein, some of the questions related to unbounded dependencies) can be directly worked on in UCG, though the first impression is that this model is less specified than GPSG. In any case, it is interesting that the proposed NP à analysis (see above p. 28 and p. 35) of some French 'prepositional' phrases has a straight UCG counterpart (see J. Calder, E. Klein and H. Zeevat Problems of Dialogue Parsing, p. 42-42).We feel that perhaps the preceding report can turn out to be an operative testing-bench for comparing in French GPSG with UCG.
